

PAPER • OPEN ACCESS

## MONKES: a fast neoclassical code for the evaluation of monoenergetic transport coefficients in stellarator plasmas

To cite this article: F.J. Escoto *et al* 2024 *Nucl. Fusion* **64** 076030

View the [article online](#) for updates and enhancements.

### You may also like

- [Fission, critical mass and safety—a historical review](#)  
Geoff Meggit
- [Controlling the nano-deformation of polymer by a reversible photo-cross-linking reaction](#)  
Dan-Thuy Van-Pham, Minh Tri Nguyen, Ken Ohdomari et al.
- [Evaluation and Implementation of Convolutional Neural Networks in Image Recognition](#)  
Zheng Yi

# MONKES: a fast neoclassical code for the evaluation of monoenergetic transport coefficients in stellarator plasmas

F.J. Escoto<sup>1,\*</sup>, J.L. Velasco<sup>1</sup>, I. Calvo<sup>1</sup>, M. Landreman<sup>2</sup> and F.I. Parra<sup>3</sup>

<sup>1</sup> Laboratorio Nacional de Fusión, CIEMAT, 28040 Madrid, Spain

<sup>2</sup> University of Maryland, College Park, MD 20742, United States of America

<sup>3</sup> Princeton Plasma Physics Laboratory, Princeton, NJ 08540, United States of America

E-mail: [fjavier.escoto@ciemat.es](mailto:fjavier.escoto@ciemat.es)

Received 19 December 2023, revised 8 March 2024

Accepted for publication 17 April 2024

Published 3 June 2024



## Abstract

MONKES is a new neoclassical code for the evaluation of monoenergetic transport coefficients in stellarators. By means of a convergence study and benchmarks with other codes, it is shown that MONKES is accurate and efficient. The combination of spectral discretization in spatial and velocity coordinates with block sparsity allows MONKES to compute monoenergetic coefficients at low collisionality, in a single core, in approximately one minute. MONKES is sufficiently fast to be integrated into stellarator optimization codes for direct optimization of the bootstrap current and to be included in predictive transport suites. The code and data from this paper are available at <https://github.com/JavierEscoto/MONKES/>.

Supplementary material for this article is available [online](#)

Keywords: stellarator optimization, neoclassical transport, bootstrap current

(Some figures may appear in colour only in the online journal)

## 1. Introduction

Stellarators are an attractive alternative to tokamaks as future fusion reactors. While tokamaks require a large toroidal current to generate part of the magnetic field, in stellarators the field is produced entirely by external magnets. As a consequence, stellarators avoid current-induced instabilities and facilitate steady-state operation. These advantages come at the

expense of making the magnetic field three-dimensional. In tokamaks, axisymmetry guarantees that the radial displacement that charged particles experience along their collisionless orbits averages to zero. Therefore, in the absence of collisions, all charged particles are confined. However, in a generic stellarator the orbit-averaged radial drift velocity does not vanish for trapped particles and they quickly drift out of the device. The combination of a non zero orbit-averaged radial drift and a small collision frequency (reactor-relevant fusion plasmas are weakly collisional in the core) produces, for a generic stellarator, intolerably large levels of neoclassical transport.

Hence, stellarator magnetic fields must be carefully designed in order to display good confinement properties. This process of tailoring of the magnetic field is called stellarator optimization. The objective of neoclassical optimization is to

\* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

obtain a stellarator with levels of neoclassical losses equivalent or lower to those in an axisymmetric device. Stellarator magnetic fields in which the orbit-averaged radial magnetic drift is zero for all particles are called omnigenous [1]. Thus, the goal of neoclassical optimization is to obtain magnetic fields which are close to omnogeneity. However, addressing only radial transport in the optimization process is not sufficient. In toroidal plasmas, the parallel flow of electrons and the rest of species is not, in general, balanced. This mismatch produces a net parallel current at each flux surface which, through Ampère's law, modifies the magnetic field  $\mathbf{B}$ . When the current is generated by a combination of neoclassical mechanisms and non-zero plasma profile gradients, we speak of bootstrap current. The bootstrap current and its effect on the magnetic configuration must be taken into account in the design of optimized stellarator magnetic fields.

Two different subclasses of omnigenous stellarators have drawn particular attention: quasi-isodynamic (QI) and quasi-symmetric (QS) stellarators. QI configurations are omnigenous configurations in which the curves of constant magnetic field strength  $B := |\mathbf{B}|$  on a flux surface close poloidally. This additional property has an important implication: exactly QI stellarators produce zero bootstrap current at low collisionality [2, 3]. Thanks to this feature, QI stellarators can control plasma-wall interaction by means of a divertor relying on a specific structure of islands, which could not be realized in the presence of large toroidal currents. The Wendelstein 7-X (W7-X) experiment was designed to be close to QI and demonstrates that theoretically based stellarator optimization can be applied to construct a device with much better confinement properties than any previous stellarator [4]. Moreover, the bootstrap current produced in W7-X plasmas is smaller than in non-optimized machines [5]. However, despite its success, there is still room for improvement. The two main configurations of W7-X, the KJM (or so-called 'high mirror') and the EIM (also known as 'standard') are not optimized for simultaneously having low levels of radial and parallel neoclassical transport [4, 6]: While W7-X EIM has small radial transport, it has intolerably large bootstrap current. Conversely, W7-X KJM displays small bootstrap current but larger levels of radial transport. Consequently, optimization of QI stellarators is a very active branch of research and, recently, much effort has been put in pushing forward the design and construction of QI stellarators [7–11].

The QS subclass of omnigenous configurations is attractive as the neoclassical properties of such magnetic fields are isomorphic to those in a tokamak [12, 13]. Recently, it has been shown that it is possible to design QS magnetic fields with extremely low neoclassical losses [14]. In contrast to QI configurations, QS stellarators are expected to have a substantial bootstrap current<sup>4</sup> and its effect must be taken into account [16]. Examples of this subclass are the Helically Symmetric eXperiment (HSX) [17], the Chinese First

Quasi-Axisymmetric Stellarator [18] or the design of the National Compact Stellarator Experiment (NCSX) [19].

Typically, at each iteration of the optimization process a large number ( $\sim 10^2$ ) of magnetic configurations are generated. Therefore, in order to neoclassically optimize magnetic fields, it is required to be able to evaluate fast the neoclassical properties of each configuration. Due to this requirement, neoclassical properties are typically addressed indirectly. For instance, one can tailor the variation of the magnetic field strength  $B$  on the flux-surface so that it nearly fulfils quasi-isodinamicity: the isolines of  $B$  can be forced to close poloidally and the variance of the extrema of  $B$  along field lines can be minimized.

A different approach relies on figures of merit, which are easy to calculate, for specific collisional regimes. For the  $1/\nu$  regime, the code NEO [20] computes the effective ripple  $\epsilon_{\text{eff}}$ , which encapsulates the dependence of radial neoclassical transport on the magnetic configuration. For transport within the flux surface, there exist long mean free path formulae for parallel flow and bootstrap current [21–23]. Although they can be computed very fast and capture some qualitative behaviour, these formulae are plagued with noise due to resonances in rational surfaces and, even with smoothing ad-hoc techniques, they are not accurate [16]. This lack of accuracy limits their application for optimization purposes. During the optimization process, an accurate calculation of the bootstrap current is required to account for its effect (e.g. for optimizing QS stellarators) or to keep it sufficiently small (when optimizing for quasi-isodinamicity).

Recent developments allow direct optimization of radial neoclassical transport. Based on previous derivations [24, 25], the code KNOSOS [26, 27] solves very fast an orbit-averaged drift-kinetic equation that is accurate for low collisionality regimes. KNOSOS is included in the stellarator optimization suite STELLOPT [28] and in the predictive transport frameworks TANGO [29] and TRINITY [30]. However, the orbit-averaged equations solved by KNOSOS only describe radial transport at low collisionalities.

In this work we present MONKES (MONoenergetic Kinetic equation Solver), a new neoclassical code conceived to satisfy the necessity of fast and accurate calculations of the bootstrap current for stellarator optimization. Specifically, MONKES makes it possible to compute the monoenergetic coefficients  $\hat{D}_{ij}$  where  $i, j \in \{1, 2, 3\}$  (their precise definition is given in section 2). These nine coefficients encapsulate neoclassical transport across and within flux surfaces. The parallel flow of each species can be calculated in terms of the coefficients  $\hat{D}_{3j}$  [31–34]. In the absence of externally applied loop voltage, the bootstrap current is driven by the radial electric field and gradients of density and temperature. The so-called bootstrap current coefficient  $\hat{D}_{31}$  is the one that relates the parallel flow to these gradients. The six remaining coefficients  $\hat{D}_{ij}$  for  $i \in \{1, 2\}$  allow to compute the flux of particles and heat across the flux surface.

MONKES also computes fast the radial transport coefficients. Although at low collisionality it is not as fast as the orbit-averaged code KNOSOS, MONKES can compute the transition from the  $1/\nu$  and  $\sqrt{\nu}\cdot\nu$  regimes to the plateau regime or

<sup>4</sup> With the exception of the quasi-poloidally symmetric magnetic field, which lies at the intersection of QI and QS configurations. However, quasi-poloidal symmetry is impossible to achieve near the magnetic axis, see e.g. [15].

the banana regime. The plateau regime may be relevant close to the edge, while the banana regime may be necessary for stellarators very close to perfect omnigenerity. Apart from optimization, MONKES can find many other applications. For instance, it can be used for the analysis of experimental discharges or also be included in predictive transport frameworks.

This paper is organized as follows: in section 2, we introduce the drift-kinetic equation solved by MONKES and the transport coefficients that it computes. In section 3, we explain the algorithm used to solve the drift-kinetic equation and its implementation. In section 4, by means of a convergence study, we demonstrate that MONKES can be used to compute accurate monoenergetic coefficients at low collisionality very fast for the  $1/\nu$  and  $\sqrt{\nu}/\nu$  regimes [25]. In order to show this, MONKES results are also benchmarked against DKES [35, 36] and, when necessary, against SFINCS [37]. Finally, in section 5 we summarize the results and discuss future lines of work.

## 2. Drift-kinetic equation and transport coefficients

MONKES solves the drift-kinetic equation

$$(\nu\xi\mathbf{b} + \mathbf{v}_E) \cdot \nabla h_a + \nu\nabla \cdot \mathbf{b} \frac{(1-\xi^2)}{2} \frac{\partial h_a}{\partial \xi} - \nu^a \mathcal{L} h_a = S_a, \quad (1)$$

where  $\mathbf{b} := \mathbf{B}/B$  is the unit vector tangent to magnetic field lines and we have employed as velocity coordinates the cosine of the pitch-angle  $\xi := \mathbf{v} \cdot \mathbf{b} / |\mathbf{v}|$  and the magnitude of the velocity  $\nu := |\mathbf{v}|$ .

We assume that the magnetic configuration has nested flux-surfaces. We denote by  $\psi \in [0, \psi_{\text{lcs}}]$  a radial coordinate that labels flux-surfaces, where  $\psi_{\text{lcs}}$  denotes the label of the last closed flux-surface. In equation (1),  $h_a$  is the non-adiabatic component of the deviation of the distribution function from a local Maxwellian for a plasma species  $a$

$$f_{\text{Ma}}(\psi, \nu) := n_a(\psi) \pi^{-3/2} v_{ta}^{-3}(\psi) \exp\left(-\frac{\nu^2}{v_{ta}^2(\psi)}\right). \quad (2)$$

Here,  $n_a$  is the density of species  $a$ ,  $v_{ta} := \sqrt{2T_a/m_a}$  is its thermal velocity,  $T_a$  its temperature (in energy units) and  $m_a$  its mass.

For the convective term in equation (1)

$$\mathbf{v}_E := \frac{\mathbf{E}_0 \times \mathbf{B}}{\langle \mathbf{B}^2 \rangle} = -\frac{E_\psi}{\langle \mathbf{B}^2 \rangle} \mathbf{B} \times \nabla \psi \quad (3)$$

denotes the incompressible  $\mathbf{E} \times \mathbf{B}$  drift approximation [25] and  $\mathbf{E}_0 = E_\psi(\psi)\nabla\psi$  is the electrostatic piece of the electric field  $\mathbf{E}$  perpendicular to the flux-surface. The symbol  $\langle \dots \rangle$  stands for the flux-surface average operation. Denoting by  $V(\psi)$  the volume enclosed by the flux-surface labelled by  $\psi$ , the flux-surface average of a function  $f$  can be defined as the limit

$$\langle f \rangle := \lim_{\delta\psi \rightarrow 0} \frac{\int_{V(\psi+\delta\psi)} f d^3r - \int_{V(\psi)} f d^3r}{V(\psi+\delta\psi) - V(\psi)}, \quad (4)$$

where  $d^3r$  is the spatial volume form.

We denote the Lorentz pitch-angle scattering operator by  $\mathcal{L}$ , which in coordinates  $(\xi, \nu)$  takes the form

$$\mathcal{L} := \frac{1}{2} \frac{\partial}{\partial \xi} \left( (1-\xi^2) \frac{\partial}{\partial \xi} \right). \quad (5)$$

In the collision operator,  $\nu^a(\nu) = \sum_b \nu^{ab}(\nu)$  and

$$\nu^{ab}(\nu) := \frac{4\pi n_b e_a^2 e_b^2}{m_a^2 v_{ta}^3} \log \Lambda_{ab} \frac{\text{erf}(\nu/v_{tb}) - G(\nu/v_{tb})}{\nu^3/v_{ta}^3} \quad (6)$$

stands for the pitch-angle collision frequency between species  $a$  and  $b$ . We denote the respective charges of each species by  $e_a$  and  $e_b$ , the Chandrasekhar function by  $G(x) = [\text{erf}(x) - (2x/\sqrt{\pi}) \exp(-x^2)]/(2x^2)$ ,  $\text{erf}(x)$  is the error function and  $\log \Lambda_{ab}$  is the Coulomb logarithm [38].

On the right-hand-side of equation (1)

$$S_a := -\mathbf{v}_{\text{ma}} \cdot \nabla \psi \left( A_{1a} + \frac{\nu^2}{v_{ta}^2} A_{2a} \right) f_{\text{Ma}} + \frac{B}{B_0} \nu \xi A_{3a} f_{\text{Ma}} \quad (7)$$

is the source term,

$$\mathbf{v}_{\text{ma}} \cdot \nabla \psi = -\frac{B\nu^2}{\Omega_a} \frac{1+\xi^2}{2B^3} \mathbf{B} \times \nabla \psi \cdot \nabla B \quad (8)$$

is the expression of the radial magnetic drift assuming ideal magnetohydrodynamical equilibrium,  $\Omega_a = e_a B / m_a$  is the gyrofrequency of species  $a$  and the flux-functions

$$A_{1a}(\psi) := \frac{d \ln n_a}{d \psi} - \frac{3}{2} \frac{d \ln T_a}{d \psi} - \frac{e_a E_\psi}{T_a}, \quad (9)$$

$$A_{2a}(\psi) := \frac{d \ln T_a}{d \psi}, \quad (10)$$

$$A_{3a}(\psi) := \frac{e_a B_0}{T_a} \frac{\langle \mathbf{E} \cdot \mathbf{B} \rangle}{\langle \mathbf{B}^2 \rangle} \quad (11)$$

are the so-called thermodynamical forces. Here,  $B_0(\psi)$  is a reference value for the magnetic field strength on the flux-surface (its explicit definition is given in section 3).

Mathematically speaking, there are still two additional conditions to completely determine the solution to equation (1). First, equation (1) must be solved imposing regularity conditions at  $\xi = \pm 1$

$$\left. \left( (1-\xi^2) \frac{\partial h_a}{\partial \xi} \right) \right|_{\xi=\pm 1} = 0. \quad (12)$$

Second, as the differential operator on the left-hand-side of equation (1) has a non trivial kernel, the solution to equation (1) is determined up to an additive function  $g(\psi, \nu)$ . This function is unimportant as it does not contribute to the neoclassical transport quantities of interest. Nevertheless, in order to have a unique solution to the drift-kinetic equation, it must be fixed by imposing an appropriate additional constraint. We will select this free function (for fixed  $(\psi, \nu)$ ) by imposing

$$\left\langle \int_{-1}^1 h_a d\xi \right\rangle = C, \quad (13)$$

for some  $C \in \mathbb{R}$ . We will discuss this further in section 3.

The drift-kinetic equation (1) is the one solved by the standard neoclassical code DKES [35, 36] using a variational principle. Although the main feature of the code SFINCS [37] is to solve a more complete neoclassical drift-kinetic equation, it can also solve equation (1).

Taking the moments  $\{\mathbf{v}_{\text{ma}} \cdot \nabla \psi, (v^2/v_{\text{ta}}^2)\mathbf{v}_{\text{ma}} \cdot \nabla \psi, v\xi B/B_0\}$  of  $h_a$  and then the flux-surface average yields, respectively, the radial particle flux, the radial heat flux and the parallel flow

$$\langle \mathbf{T}_a \cdot \nabla \psi \rangle := \left\langle \int \mathbf{v}_{\text{ma}} \cdot \nabla \psi \ h_a d^3v \right\rangle, \quad (14)$$

$$\left\langle \frac{\mathbf{Q}_a \cdot \nabla \psi}{T_a} \right\rangle := \left\langle \int \frac{v^2}{v_{\text{ta}}^2} \mathbf{v}_{\text{ma}} \cdot \nabla \psi \ h_a d^3v \right\rangle, \quad (15)$$

$$\frac{\langle n_a \mathbf{V}_a \cdot \mathbf{B} \rangle}{B_0} := \left\langle \frac{B}{B_0} \int v\xi \ h_a d^3v \right\rangle. \quad (16)$$

It is a common practice for linear drift-kinetic equations (e.g. [6, 35, 37]) to apply superposition and split  $h_a$  into several additive terms. As in the drift-kinetic equation (1) there are no derivatives or integrals along  $\psi$  nor  $v$ , it is convenient to use the splitting

$$h_a = f_{\text{Ma}} \left[ \frac{Bv}{\Omega_a} \left( A_{1a} f_1 + A_{2a} \frac{v^2}{v_{\text{ta}}^2} f_2 \right) + A_{3a} f_3 \right]. \quad (17)$$

The splitting is chosen so that the functions  $\{f_j\}_{j=1}^3$  are solutions to

$$\begin{aligned} & \xi \mathbf{b} \cdot \nabla f_j + \nabla \cdot \mathbf{b} \frac{(1-\xi^2)}{2} \frac{\partial f_j}{\partial \xi} \\ & - \frac{\widehat{E}_\psi}{\langle B^2 \rangle} \mathbf{B} \times \nabla \psi \cdot \nabla f_j - \hat{\nu} \mathcal{L} f_j = s_j, \end{aligned} \quad (18)$$

for  $j = 1, 2, 3$ , where  $\hat{\nu} := \nu(v)/v$  and  $\widehat{E}_\psi := E_\psi/v$ . The source terms are defined as

$$s_1 := -\mathbf{v}_{\text{ma}} \cdot \nabla \psi \frac{\Omega_a}{Bv^2}, \quad s_2 := s_1, \quad s_3 := \xi \frac{B}{B_0}. \quad (19)$$

Note that each source  $s_j$  corresponds to one of the three thermodynamic forces on the right-hand side of definition (7).

The relation between  $h_a$  and  $f_j$  given by equation (17) is such that the transport quantities (14)–(16) can be written in terms of four transport coefficients which, for fixed  $(\hat{\nu}, \widehat{E}_\psi)$ , depend only on the magnetic configuration. As  $d\hat{\nu}/dv$  never vanishes, the dependence of  $f_j$  on the velocity  $v$  can be parametrized by its dependence on  $\hat{\nu}$ . Thus, for fixed  $(\hat{\nu}, \widehat{E}_\psi)$ , equation (18) is completely determined by the magnetic configuration. Hence, its unique solutions  $f_j$  that satisfy conditions (12) and (13) are also completely determined by the magnetic configuration. The assumptions that lead to  $\psi$  and  $v$  appearing as parameters in the drift-kinetic equation (1) comprise the so-called local monoenergetic approximation to neoclassical transport (see e.g. [39]).

Using splitting (17) we can write the transport quantities (14)–(16) in terms of the Onsager matrix

$$\begin{bmatrix} \langle \mathbf{T}_a \cdot \nabla \psi \rangle \\ \left\langle \frac{\mathbf{Q}_a \cdot \nabla \psi}{T_a} \right\rangle \\ \left\langle \frac{n_a \mathbf{V}_a \cdot \mathbf{B}}{B_0} \right\rangle \end{bmatrix} = \begin{bmatrix} L_{11a} & L_{12a} & L_{13a} \\ L_{21a} & L_{22a} & L_{23a} \\ L_{31a} & L_{32a} & L_{33a} \end{bmatrix} \begin{bmatrix} A_{1a} \\ A_{2a} \\ A_{3a} \end{bmatrix}. \quad (20)$$

Here, we have defined the thermal transport coefficients as

$$L_{ija} := \int_0^\infty 2\pi v^2 f_{\text{Ma}} w_i w_j D_{ija} dv, \quad (21)$$

where  $w_1 = w_3 = 1$ ,  $w_2 = v^2/v_{\text{ta}}^2$  and we have used that  $\int g d^3v = 2\pi \int_0^\infty \int_{-1}^1 g v^2 d\xi dv$  for any integrable function  $g(\xi, v)$ . The quantities  $D_{ija}$  are defined as

$$D_{ija} := -\frac{B^2 v^3}{\Omega_a^2} \widehat{D}_{ij}, \quad i, j \in \{1, 2\}, \quad (22)$$

$$D_{i3a} := -\frac{Bv^2}{\Omega_a} \widehat{D}_{i3}, \quad i \in \{1, 2\}, \quad (23)$$

$$D_{3ja} := \frac{Bv^2}{\Omega_a} \widehat{D}_{3j}, \quad j \in \{1, 2\}, \quad (24)$$

$$D_{33a} := v \widehat{D}_{33}, \quad (25)$$

where

$$\widehat{D}_{ij}(\psi, v) := \left\langle \int_{-1}^1 s_i f_j d\xi \right\rangle, \quad i, j \in \{1, 2, 3\} \quad (26)$$

are the monoenergetic geometric coefficients. Note that (unlike  $D_{ija}$ ) the monoenergetic geometric coefficients  $\widehat{D}_{ij}$  do not depend on the species for fixed  $\hat{\nu}$  (however the correspondent value of  $v$  associated to each  $\hat{\nu}$  varies between species) and depend only on the magnetic geometry. In general, four independent monoenergetic geometric coefficients can be obtained by solving (18):  $\widehat{D}_{11}$ ,  $\widehat{D}_{13}$ ,  $\widehat{D}_{31}$  and  $\widehat{D}_{33}$ . However, when the magnetic field possesses stellarator symmetry [40] or there is no radial electric field, Onsager symmetry implies  $\widehat{D}_{13} = -\widehat{D}_{31}$  [36] making only three of them independent (for further details see appendix A). Hence, obtaining the transport coefficients for all species requires to solve (18) for two different source terms  $s_1$  and  $s_3$ . The algorithm for solving equation (18) is described in section 3.

Finally, we briefly comment on the validity of the coefficients provided by equation (18) for the calculation of the bootstrap current. The pitch-angle scattering collision operator used in equation (1) lacks parallel momentum conservation. Besides, the pitch-angle scattering operator is not adequate for calculating parallel flow of electrons, which is a quantity required to compute the bootstrap current. Hence, in principle, the parallel transport directly predicted by equation (1) is not correct. Fortunately, there exist techniques [31–34] to calculate the radial and parallel transport associated to more accurate momentum conserving collision operators by just solving the simplified drift-kinetic equation (18). This has been done successfully in the past by the code PENTA [32, 41], using the

results of DKES. Nevertheless, the momentum restoring technique is not needed for minimizing the bootstrap current. In the method presented in section V of [34], when there is no net parallel inductive electric field (i.e.  $A_{3a} = 0$ ), the parallel flow with the correct collision operator for any species vanishes when two integrals in  $v$  of  $\hat{D}_{31}$  vanish. Thus, minimizing  $\hat{D}_{31}$  translates in a minimization of the parallel flows of all species involved in the bootstrap current calculation, and therefore of this current.

### 3. Numerical method

In this section we describe the algorithm to numerically solve the drift-kinetic equation (18) and its implementation. The algorithm, based on the tridiagonal representation of the drift-kinetic equation, emerges naturally when the velocity coordinate  $\xi$  is discretized using a Legendre spectral method.

First, in section 3.1 we will present the algorithm in a formal way. We will use (right-handed) Boozer coordinates<sup>5</sup>  $(\psi, \theta, \zeta) \in [0, \psi_{\text{lfcs}}] \times [0, 2\pi] \times [0, 2\pi/N_p]$ . The integer  $N_p \geq 1$  denotes the number of toroidal periods of the device. The radial coordinate is selected so that  $2\pi\psi$  is the toroidal flux of the magnetic field and  $\theta, \zeta$  are respectively the poloidal and toroidal (in a single period) angles. In these coordinates, the magnetic field can be written as

$$\begin{aligned} \mathbf{B} &= \nabla\psi \times \nabla\theta - \iota(\psi) \nabla\psi \times \nabla\zeta \\ &= B_\psi(\psi, \theta, \zeta) \nabla\psi + B_\theta(\psi) \nabla\theta + B_\zeta(\psi) \nabla\zeta, \end{aligned} \quad (27)$$

and the Jacobian of the transformation reads

$$\sqrt{g}(\psi, \theta, \zeta) := (\nabla\psi \times \nabla\theta \cdot \nabla\zeta)^{-1} = \frac{B_\zeta + \iota B_\theta}{B^2}, \quad (28)$$

where  $\iota := \mathbf{B} \cdot \nabla\theta / \mathbf{B} \cdot \nabla\zeta$  is the rotational transform. The flux-surface average operation (4) is written in Boozer angles as

$$\langle f \rangle = \left( \frac{dV}{d\psi} \right)^{-1} \oint \oint f \sqrt{g} d\theta d\zeta. \quad (29)$$

We define the reference value for the magnetic field strength  $B_0$  introduced in definitions (7) and (11) as the  $(0, 0)$  Fourier mode of the magnetic field strength. Namely,

$$B_0(\psi) := \frac{N_p}{4\pi^2} \oint \oint B(\psi, \theta, \zeta) d\theta d\zeta. \quad (30)$$

Using (27) and (28), the spatial differential operators present in the drift-kinetic equation (18) can be expressed in these coordinates as

$$\mathbf{b} \cdot \nabla = \frac{\mathbf{B}}{B_\zeta + \iota B_\theta} \left( \iota \frac{\partial}{\partial\theta} + \frac{\partial}{\partial\zeta} \right), \quad (31)$$

<sup>5</sup> Even though we use Boozer coordinates, we want to stress out that the algorithm presented in section 3.1 is valid for any set of spatial coordinates in which  $\psi$  labels flux-surfaces and the two remaining coordinates parametrize the flux-surface.

$$\mathbf{B} \times \nabla\psi \cdot \nabla = \frac{B^2}{B_\zeta + \iota B_\theta} \left( B_\zeta \frac{\partial}{\partial\theta} - B_\theta \frac{\partial}{\partial\zeta} \right). \quad (32)$$

After the explanation of the algorithm, in section 3.2 its implementation in MONKES is described. In order to ease the notation, in sections 3.1 and 3.2 we drop when possible the subscript  $j$  that labels every different source term. Also, as  $\psi$  and  $v$  act as mere parameters, we will omit their dependence and functions of these two variables will be referred to as constants.

#### 3.1. Legendre polynomial expansion

The algorithm is based on the approximate representation of the distribution function  $f$  by a truncated Legendre series. We will search for approximate solutions to equation (18) of the form

$$f(\theta, \zeta, \xi) = \sum_{k=0}^{N_\xi} f^{(k)}(\theta, \zeta) P_k(\xi), \quad (33)$$

where  $f^{(k)} = \langle f, P_k \rangle_{\mathcal{L}} / \langle P_k, P_k \rangle_{\mathcal{L}}$  is the  $k$ -th Legendre mode of  $f(\theta, \zeta, \xi)$  (see appendix B) and  $N_\xi$  is an integer greater or equal to 1. As mentioned in appendix B, the expansion in Legendre polynomials (33) ensures that the regularity conditions (12) are satisfied. Of course, in general, the exact solution to equation (18) does not have a finite Legendre spectrum, but taking  $N_\xi$  sufficiently high in expansion (33) yields an approximate solution to the desired degree of accuracy (in infinite precision arithmetic).

In appendix B we derive explicitly the projection of each term of the drift-kinetic equation (18) onto the Legendre basis when the representation (33) is used. When doing so, we obtain that the Legendre modes of the drift-kinetic equation have the tridiagonal representation

$$L_k f^{(k-1)} + D_k f^{(k)} + U_k f^{(k+1)} = s^{(k)}, \quad (34)$$

for  $k = 0, 1, \dots, N_\xi$ , where we have defined for convenience  $f^{(-1)} := 0$  and from expansion (33) it is clear that  $f^{(N_\xi+1)} = 0$ . Analogously to (33) the source term is expanded as  $s = \sum_{k=0}^{N_\xi} s^{(k)} P_k$ . For the sources given by (19) this expansion is exact when  $N_\xi \geq 2$  as  $s_j^{(k)} = 0$  for  $k \geq 3$ . The spatial differential operators read

$$L_k = \frac{k}{2k-1} \left( \mathbf{b} \cdot \nabla + \frac{k-1}{2} \mathbf{b} \cdot \nabla \ln B \right), \quad (35)$$

$$D_k = -\frac{\hat{E}_\psi}{\langle B^2 \rangle} \mathbf{B} \times \nabla\psi \cdot \nabla + \frac{k(k+1)}{2} \hat{\nu}, \quad (36)$$

$$U_k = \frac{k+1}{2k+3} \left( \mathbf{b} \cdot \nabla - \frac{k+2}{2} \mathbf{b} \cdot \nabla \ln B \right). \quad (37)$$

Thanks to its tridiagonal structure, the system of equations (34) can be inverted using the standard Gaussian elimination algorithm for block tridiagonal matrices.

Before introducing the algorithm we will explain how to fix the free constant of the solution to equation (34) so that it can

be inverted. Note that the aforementioned kernel of the drift-kinetic equation translates in the fact that  $f^{(0)}$  is not completely determined from equation (34). To prove this, we inspect the modes  $k=0$  and  $k=1$  of equation (34), which are the ones that involve  $f^{(0)}$ . From expression (32) we can deduce that the term  $D_0 f^{(0)} + U_0 f^{(1)}$  is invariant if we add to  $f^{(0)}$  any function of  $B_\theta \theta + B_\zeta \zeta$ . For  $\widehat{E}_\psi \neq 0$ , functions of  $B_\theta \theta + B_\zeta \zeta$  lie on the kernel of  $\mathbf{B} \times \nabla \psi \cdot \nabla$  and for  $\widehat{E}_\psi = 0$ ,  $D_0$  is identically zero. Besides, the term  $L_1 f^{(0)} + D_1 f^{(1)} + U_1 f^{(2)}$  remains invariant if we add to  $f^{(0)}$  any function of  $\theta - \nu \zeta$  (the kernel of  $L_1 = \mathbf{b} \cdot \nabla$  consists of these functions). For ergodic flux-surfaces, the only continuous functions on the torus that belong to the kernel of  $L_1$  are constants. Thus, equation (34) is unaltered when we add to  $f^{(0)}$  any constant (a function that belongs simultaneously to the kernels of  $\mathbf{B} \times \nabla \psi \cdot \nabla$  and  $\mathbf{b} \cdot \nabla$ ). A constraint equivalent to condition (13) is to fix the value of the 0-th Legendre mode of the distribution function at a single point of the flux-surface. For example,

$$f^{(0)}(0,0) = 0, \quad (38)$$

which implicitly fixes the value of the constant  $C$  in (13). With this condition, equation (34) has a unique solution and its left-hand-side can be inverted to solve for  $f^{(k)}$  in two scenarios: when the flux-surface is ergodic and in rational surfaces when  $\widehat{E}_\psi \neq 0$  (further details on its invertibility are given in appendix C). Note that, as expansion (33) is finite and representation (34) is non diagonal, the functions  $f^{(k)}$  obtained from inverting (34) constrained by (38) are approximations to the first  $N_\xi + 1$  Legendre modes of the exact solution to (18) satisfying (13) (further details at the end of appendix B).

The algorithm for solving the truncated drift-kinetic equation (34) consists of two steps.

### (i) Forward elimination

Starting from  $\Delta_{N_\xi} = D_{N_\xi}$  and  $\sigma^{(N_\xi)} = s^{(N_\xi)}$  we can obtain recursively the operators

$$\Delta_k = D_k - U_k \Delta_{k+1}^{-1} L_{k+1}, \quad (39)$$

and the sources

$$\sigma^{(k)} = s^{(k)} - U_k \Delta_{k+1}^{-1} \sigma^{(k+1)}, \quad (40)$$

for  $k = N_\xi - 1, N_\xi - 2, \dots, 0$  (in this order). Equations (39) and (40) define the forward elimination. With this procedure we can transform equation (34) to the equivalent system

$$L_k f^{(k-1)} + \Delta_k f^{(k)} = \sigma^{(k)}, \quad (41)$$

for  $k = 0, 1, \dots, N_\xi$ . Note that this process corresponds to perform formal Gaussian elimination over

$$\left[ \begin{array}{ccc|c} L_k & D_k & U_k & s^{(k)} \\ 0 & L_{k+1} & \Delta_{k+1} & \sigma^{(k+1)} \end{array} \right], \quad (42)$$

to eliminate  $U_k$  in the first row.

### (ii) Backward substitution

Once we have the system of equations in the form (41) it is immediate to solve recursively

$$f^{(k)} = \Delta_k^{-1} (\sigma^{(k)} - L_k f^{(k-1)}), \quad (43)$$

for  $k = 0, 1, \dots, N_\xi$  (in this order). Here,  $\Delta_0^{-1} \sigma^{(0)}$  denotes the unique solution to  $\Delta_0 f^{(0)} = \sigma^{(0)}$  that satisfies (38). As  $L_1 = \mathbf{b} \cdot \nabla$ , using expression (31), it is clear from equation (43) that the integration constant does not affect the value of  $f^{(1)}$ .

We can apply this algorithm to solve equation (18) for  $f_1$ ,  $f_2$  and  $f_3$  in order to compute approximations to the transport coefficients. In terms of the Legendre modes of  $f_1$ ,  $f_2$  and  $f_3$ , the monoenergetic geometric coefficients from definition (26) read

$$\widehat{D}_{11} = 2 \left\langle s_1^{(0)} f_1^{(0)} \right\rangle + \frac{2}{5} \left\langle s_1^{(2)} f_1^{(2)} \right\rangle, \quad (44)$$

$$\widehat{D}_{31} = \frac{2}{3} \left\langle \frac{\mathbf{B}}{B_0} f_1^{(1)} \right\rangle, \quad (45)$$

$$\widehat{D}_{13} = 2 \left\langle s_1^{(0)} f_3^{(0)} \right\rangle + \frac{2}{5} \left\langle s_1^{(2)} f_3^{(2)} \right\rangle, \quad (46)$$

$$\widehat{D}_{33} = \frac{2}{3} \left\langle \frac{\mathbf{B}}{B_0} f_3^{(1)} \right\rangle, \quad (47)$$

where  $3s_1^{(0)}/2 = 3s_1^{(2)} = \mathbf{B} \times \nabla \psi \cdot \nabla B / B^3$ . Note from expressions (44)–(47) that, in order to compute the monoenergetic geometric coefficients  $\widehat{D}_{ij}$ , we only need to calculate the Legendre modes  $k = 0, 1, 2$  of the solution and we can stop the backward substitution (43) at  $k = 2$ . In the next subsection we will explain how MONKES solves equation (34) using this algorithm.

### 3.2. Spatial discretization and algorithm implementation

The algorithm described above allows, in principle, to compute the exact solution to the truncated drift-kinetic equation (34) which is an approximate solution to (18). However, to our knowledge, it is not possible to give an exact expression for the operator  $\Delta_k^{-1}$  except for  $k = N_\xi \geq 1$ . Instead, we are forced to compute an approximate solution to (34). In order to obtain an approximate solution of equation (34) we assume that each  $f^{(k)}$  has a finite Fourier spectrum so that it can be expressed as

$$f^{(k)}(\theta, \zeta) = \mathbf{I}(\theta, \zeta) \cdot \mathbf{f}^{(k)}, \quad (48)$$

where the Fourier interpolant row vector map  $\mathbf{I}(\theta, \zeta)$  is defined at appendix D and the column vector  $\mathbf{f}^{(k)} \in \mathbb{R}^{N_{fs}}$  contains  $f^{(k)}$  evaluated at the equispaced grid points

$$\theta_i = 2\pi i / N_\theta, \quad i = 0, 1, \dots, N_\theta - 1, \quad (49)$$

$$\zeta_j = 2\pi j / (N_\zeta N_p), \quad j = 0, 1, \dots, N_\zeta - 1. \quad (50)$$

Here,  $N_{fs} := N_\theta N_\zeta$  is the number of points in which we discretize the flux-surface being  $N_\theta$  and  $N_\zeta$  respectively the number of points in which we divide the domains of  $\theta$  and  $\zeta$ . In general,

the solution to equation (34) has an infinite Fourier spectrum and cannot exactly be written as (48) but, taking sufficiently large values of  $N_\theta$  and  $N_\zeta$ , we can approximate the solution to equation (34) to arbitrary degree of accuracy (in infinite precision arithmetic). As explained in appendix D, introducing the Fourier interpolant (48) in equation (34) and then evaluating the result at the grid points provides a system of  $N_{\text{fs}} \times (N_\xi + 1)$  equations which can be solved for  $\{\mathbf{f}^{(k)}\}_{k=0}^{N_\xi}$ . This system of equations is obtained by substituting the operators  $L_k, D_k, U_k$  in equation (34) by the  $N_{\text{fs}} \times N_{\text{fs}}$  matrices  $\mathbf{L}_k, \mathbf{D}_k, \mathbf{U}_k$ , defined in appendix D. Thus, we discretize (34) as

$$\mathbf{L}_k \mathbf{f}^{(k-1)} + \mathbf{D}_k \mathbf{f}^{(k)} + \mathbf{U}_k \mathbf{f}^{(k+1)} = \mathbf{s}^{(k)}, \quad (51)$$

for  $k = 0, 1, \dots, N_\xi$  where  $\mathbf{s}^{(k)} \in \mathbb{R}^{N_{\text{fs}}}$  contains  $s^{(k)}$  evaluated at the equispaced grid points. This system has a block tridiagonal structure and the algorithm presented in section 3.1 can be applied. We just have to replace in equations (39), (40) and (43) the operators and functions by their respective matrix and vector analogues, which we denote by boldface letters.

The matrix approximation to the forward elimination procedure given by equations (39) and (40) reads

$$\Delta_k = \mathbf{D}_k - \mathbf{U}_k \Delta_{k+1}^{-1} \mathbf{L}_{k+1}, \quad (52)$$

$$\boldsymbol{\sigma}^{(k)} = \mathbf{s}^{(k)} - \mathbf{U}_k \Delta_{k+1}^{-1} \boldsymbol{\sigma}^{(k+1)}, \quad (53)$$

for  $k = N_\xi - 1, N_\xi - 2, \dots, 0$  (in this order). Thus, starting from  $\Delta_{N_\xi} = \mathbf{D}_{N_\xi}$  and  $\boldsymbol{\sigma}^{(N_\xi)} = \mathbf{s}^{(N_\xi)}$  all the matrices  $\Delta_k$  and the vectors  $\boldsymbol{\sigma}^{(k)}$  are defined from equations (52) and (53). Obtaining the matrix  $\Delta_k$  directly from equation (52) requires to invert  $\Delta_{k+1}$ , perform two matrix multiplications and a subtraction of matrices. The inversion using LU factorization and each matrix multiplication require  $O(N_{\text{fs}}^3)$  operations so it is desirable to reduce the number of matrix multiplications as much as possible. We can reduce the number of matrix multiplications in determining  $\Delta_k$  to one if instead of computing  $\Delta_{k+1}^{-1}$  we solve the matrix system of equations

$$\Delta_{k+1} \mathbf{X}_{k+1} = \mathbf{L}_{k+1}, \quad (54)$$

for  $\mathbf{X}_{k+1}$  and then obtain

$$\Delta_k = \mathbf{D}_k - \mathbf{U}_k \mathbf{X}_{k+1}, \quad (55)$$

for  $k = N_\xi - 1, N_\xi - 2, \dots, 0$ . Thus, obtaining  $\Delta_k$  requires  $O(N_{\text{fs}}^3)$  operations for solving equation (54) (using LU factorization) and also  $O(N_{\text{fs}}^3)$  operations for applying (55). For computing the monoenergetic coefficients, the backward substitution step requires solving equation (41) for  $k = 0, 1$  and 2. Therefore, for  $k \leq 1$ , it is convenient to store  $\Delta_{k+1}$  in the factorized LU form obtained when equation (54) was solved for  $\mathbf{X}_{k+1}$ . The matrix  $\Delta_0$  will be factorized later, during the backward substitution step.

Similarly to what is done to obtain  $\Delta_k$ , to compute  $\boldsymbol{\sigma}^{(k)}$  we first solve

$$\Delta_{k+1} \mathbf{y} = \boldsymbol{\sigma}^{(k+1)} \quad (56)$$

for  $\mathbf{y}$  and then compute

$$\boldsymbol{\sigma}^{(k)} = \mathbf{s}^{(k)} - \mathbf{U}_k \mathbf{y}, \quad (57)$$

for  $k \geq 0$ . Recall that none of the source terms  $s_1, s_2$  and  $s_3$  defined by (19) have Legendre modes greater than 2. Specifically, equation (53) implies  $\boldsymbol{\sigma}_1^{(k)}, \boldsymbol{\sigma}_3^{(k-1)} = 0$  for  $k \geq 3$  and also  $\boldsymbol{\sigma}_1^{(2)} = \mathbf{s}_1^{(2)}, \boldsymbol{\sigma}_3^{(1)} = \mathbf{s}_3^{(1)}$ . Thus, we only have to solve equation (56) and apply (57) to obtain  $\{\boldsymbol{\sigma}_1^{(k)}\}_{k=0}^1$  and  $\boldsymbol{\sigma}_3^{(0)}$ . As  $\{\Delta_{k+1}\}_{k=0}^1$  are already LU factorized, solving equation (56) and then applying (57) requires  $O(N_{\text{fs}}^2)$  operations and its contribution to the arithmetic complexity of the algorithm is subdominant with respect to the  $O(N_{\text{fs}}^3)$  operations required to compute  $\Delta_k$ .

For the backward substitution, we first note that solving the matrix version of equation (41) to obtain  $\mathbf{f}^{(0)}$  requires  $O(N_{\text{fs}}^3)$  operations, as  $\Delta_0$  has not been LU factorized during the forward elimination. On the other hand, obtaining the remaining modes  $\{\mathbf{f}^{(k)}\}_{k=1}^2$  requires  $O(N_{\text{fs}}^2)$  operations. As the resolution of the matrix system of equations (54) and the matrix multiplication in (55) must be done  $N_\xi$  times, solving equation (51) by this method requires  $O(N_\xi N_{\text{fs}}^3)$  operations.

In what concerns to memory resources, as we are only interested in the Legendre modes 0, 1 and 2, it is not necessary to store in memory all the matrices  $\mathbf{L}_k, \mathbf{D}_k, \mathbf{U}_k$  and  $\Delta_k$ . Instead, we store solely  $\mathbf{L}_k, \mathbf{U}_k$  and  $\Delta_k$  (in LU form) for  $k = 0, 1, 2$ . For the intermediate steps we just need to use some auxiliary matrices  $\mathbf{L}, \mathbf{D}, \mathbf{U}, \Delta$  and  $X$  of size  $N_{\text{fs}}$ . This makes the amount of memory required by MONKES independent of  $N_\xi$ , being of order  $N_{\text{fs}}^2$ .

To summarize, the pseudocode of the implementation of the algorithm in MONKES is given in algorithm 1. In the first loop from  $k = N_\xi - 1$  to  $k = 0$  we construct and save only the matrices  $\{\mathbf{L}_k, \mathbf{U}_k, \Delta_k\}_{k=0}^2$ . At this point the matrices  $\{\Delta_k\}_{k=1}^2$  are factorized in LU form. In the second loop, the sources  $\{\boldsymbol{\sigma}_1^{(k)}\}_{k=0}^1$  and  $\boldsymbol{\sigma}_3^{(0)}$  are computed and saved for the backward substitution. Finally, the backward substitution step is applied. For solving  $\Delta_0 \mathbf{f}^{(0)} = \boldsymbol{\sigma}^{(0)}$  we have to perform the LU factorization of  $\Delta_0$  (just for one of the two source terms) and then solve for  $\mathbf{f}^{(0)}$ . For the remaining modes, the LU factorizations of  $\{\Delta_k\}_{k=1}^2$  are reused to solve for  $\{\mathbf{f}^{(k)}\}_{k=1}^2$ .

Once we have solved equation (51) for  $\mathbf{f}^{(0)}, \mathbf{f}^{(1)}$  and  $\mathbf{f}^{(2)}$ , the integrals of the flux-surface average operation involved in the monoenergetic coefficients (44)–(47), are conveniently computed using the trapezoidal rule, which for periodic analytic functions has geometric convergence [42]. In section 4 we will see that despite the cubic scaling in  $N_{\text{fs}}$  of the arithmetical complexity of the algorithm, it is possible to obtain fast and accurate calculations of the monoenergetic geometric coefficients at low collisionality (and in particular  $\widehat{D}_{31}$ ) in a single core. The reason behind this is that in the asymptotic relation  $O(N_{\text{fs}}^3) \sim C_{\text{alg}} N_{\text{fs}}^3$ , the constant  $C_{\text{alg}}$  is small enough to allow  $N_{\text{fs}}$  to take a sufficiently high value to capture accurately the spatial dependence of the distribution function without increasing much the wall-clock time.

**Algorithm 1.** Block tridiagonal solution algorithm implemented in MONKES.

**1. Forward elimination:**

```

 $L \leftarrow L_{N_\xi}$                                 ▷ Starting value for  $L$ 
 $\Delta \leftarrow D_{N_\xi}$                                 ▷ Starting value for  $\Delta$ 
Solve  $\Delta X = L$                                 ▷ Compute  $X_{N_\xi}$  stored in  $X$ 
for  $k = N_\xi - 1$  to 0 do
   $L \leftarrow L_k$                                 ▷ Construct  $L_k$  stored in  $L$ 
   $D \leftarrow D_k$                                 ▷ Construct  $D_k$  stored in  $D$ 
   $U \leftarrow U_k$                                 ▷ Construct  $U_k$  stored in  $U$ 
   $\Delta \leftarrow D - UX$                                 ▷ Construct  $\Delta_k$  stored in  $\Delta$ 
  if  $k > 0$ : Solve  $\Delta X = L$                                 ▷ Compute  $X_k$  stored in  $X$  for
                                                next iteration
  if  $k \leq 2$  then
    if  $k = 0$ :  $L_k \leftarrow L$                                 ▷ Save required matrices
     $U_k \leftarrow U$                                 ▷ Save  $\{L_k\}_{k=1}^2$ 
     $\Delta_k \leftarrow \Delta$                                 ▷ Save  $\{U_k\}_{k=0}^2$ 
  end if
end for

for  $k = 1$  to 0 do
  Solve  $\Delta_{k+1}y_1 = \sigma_1^{(k+1)}$ 
  if  $k = 0$ : Solve  $\Delta_{k+1}y_3 = \sigma_3^{(k+1)}$ 
   $\sigma_1^{(k)} \leftarrow s_1^{(k)} - U_k y_1$                                 ▷ Construct  $\sigma_1^{(k)}$ 
  if  $k = 0$ :  $\sigma_3^{(0)} \leftarrow -U_0 y_3$                                 ▷ Construct  $\sigma_3^{(0)}$ 
end for

```

**2. Backward substitution:**

```

Solve  $\Delta_0 f^{(0)} = \sigma^{(0)}$ 
for  $k = 1$  to 2
  Solve  $\Delta_k f^{(k)} = \sigma^{(k)} - L_k f^{(k-1)}$ 
end for

```

The algorithm is implemented in the new code MONKES, written in Fortran language. The matrix inversions and multiplications are computed using the linear algebra library LAPACK [43].

#### 4. Code performance and benchmark

In this section we will demonstrate that MONKES provides fast and accurate calculations of the monoenergetic coefficients from low ( $\hat{\nu} = 10^{-5} \text{ m}^{-1}$ ) to high collisionality<sup>6</sup> ( $\hat{\nu} = 3 \cdot 10^2 \text{ m}^{-1}$ ).

In section 4.1 we will see that for a correct calculation of the monoenergetic coefficients for  $\hat{\nu} \geq 10^{-5} \text{ m}$ ,  $N_{\text{fs}} \lesssim 2000$  and  $N_\xi \lesssim 200$  are required. In section 4.2 it is shown that for these resolutions MONKES produces fast calculations in a single processor. Finally, in section 4.3 the coefficients computed with MONKES will be benchmarked with DKES and SFINCS. As a result of the benchmarking, we will conclude that MONKES calculations are accurate.

**Table 1.** Cases considered in the convergence study of monoenergetic coefficients and values of  $(\hat{\nu}, \hat{E}_r)$ .

| Configuration | $\psi/\psi_{\text{lcf}}$ | $\hat{\nu} [\text{m}^{-1}]$ | $\hat{E}_r [\text{kV} \cdot \text{s m}^{-2}]$ |
|---------------|--------------------------|-----------------------------|---|
| W7X-EIM       | 0.200                    | $10^{-5}$                   | 0   |
| W7X-EIM       | 0.200                    | $10^{-5}$                   | $3 \cdot 10^{-4}$                             |
| W7X-KJM       | 0.204                    | $10^{-5}$                   | 0   |
| W7X-KJM       | 0.204                    | $10^{-5}$                   | $3 \cdot 10^{-4}$                             |
| CIEMAT-QI     | 0.250                    | $10^{-5}$                   | 0   |
| CIEMAT-QI     | 0.250                    | $10^{-5}$                   | $10^{-3}$                                     |

#### 4.1. Convergence of monoenergetic coefficients at low collisionality

In low collisionality regimes, convection is dominant with respect to diffusion. As equation (18) is singularly perturbed with respect to  $\hat{\nu}$ , its solution possesses internal boundary layers in  $\xi$ . These boundary layers appear at the interfaces between different classes of trapped particles. At these regions of phase space, collisions are no longer subdominant with respect to advection. Besides, at these regions, the poloidal  $E \times B$  precession from equation (18) can produce the chaotic transition of collisionless particles from one class to another due to separatrix crossing mechanisms [25, 44]. The existence of these localized regions with large  $\xi$  gradients demands a high number of Legendre modes  $N_\xi$ , explaining the difficulty to obtain fast and accurate solutions to equation (18) at low collisionality.

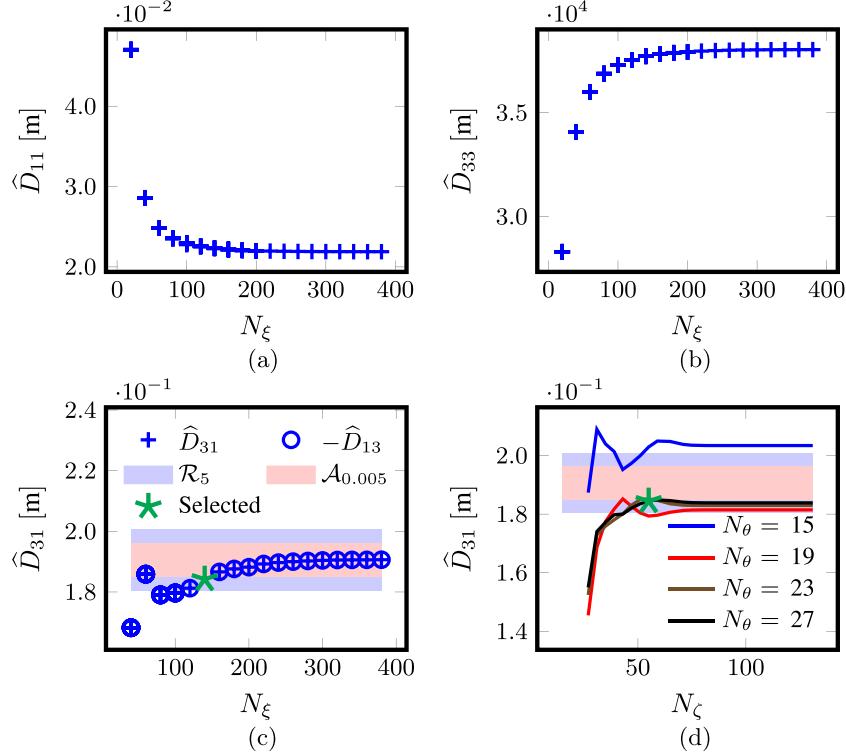
In this subsection we will select resolutions  $N_\theta$ ,  $N_\zeta$  and  $N_\xi$  for which MONKES provides accurate calculations of the monoenergetic coefficients in a wide range of collisionalities. For this, we will study how the monoenergetic coefficients computed by MONKES converge with  $N_\theta$ ,  $N_\zeta$  and  $N_\xi$  at low collisionality. From the point of view of numerical analysis, the need for large values of  $N_\xi$  is due to the lack of diffusion along  $\xi$  in equation (18). Hence, if MONKES is capable of producing fast and accurate calculations at low collisionality, it will also produce fast and accurate calculations at higher collisionalities.

For the convergence study, we select three different magnetic configurations at a single flux surface. Two of them correspond to configurations of W7-X: EIM and KJM. The third one corresponds to the new QI ‘flat mirror’ [8] configuration CIEMAT-QI [7]. The calculations are done for the  $1/\nu$  (cases with  $\hat{E}_r = 0$ ) and  $\sqrt{\nu}/\nu$  regimes [25] (cases with  $\hat{E}_r \neq 0$ ) at the low collisionality value  $\hat{\nu} = 10^{-5} \text{ m}$ . In table 1 the cases considered are listed, including their correspondent values of  $\hat{E}_r := \hat{E}_\psi d\psi/dr$ . We have denoted  $r = a\sqrt{\psi/\psi_{\text{lcf}}}$  and, in this context,  $a$  is the minor radius of the device<sup>7</sup>.

In order to select the triplets  $(N_\theta, N_\zeta, N_\xi)$  for sufficiently accurate calculations of  $\hat{D}_{31}$ , we need to specify when we will consider that a computation has converged. For each case of table 1 we will proceed in the same manner. First, we plot

<sup>6</sup> In this context ‘accurate at high collisionality’ means that the drift-kinetic equation (18) is solved accurately.

<sup>7</sup> DKES uses  $r$  as radial coordinate instead of  $\psi$ . The quantities  $\hat{\nu}$  and  $\hat{E}_r$  are denoted respectively CMUL and EFIELD in the code DKES.



**Figure 1.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for W7X-EIM at the surface labelled by  $\psi/\psi_{\text{lfcs}} = 0.200$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ .

the coefficients  $\hat{D}_{ij}$  as functions of the number of Legendre modes in a sufficiently wide interval. For each value of  $N_\xi$ , the selected spatial resolutions  $N_\theta$  and  $N_\zeta$  are large enough so that increasing them varies the monoenergetic coefficients in less than a 1%. We will say that these calculations are ‘spatially converged’. Since, typically, the most difficult coefficient to calculate is the bootstrap current coefficient, we will select the resolutions so that  $\hat{D}_{31}$  is accurately computed. From the curve of (spatially converged)  $\hat{D}_{31}$  as a function of  $N_\xi$  we define our converged reference value, which we denote by  $\hat{D}_{31}^r$ , as the converged calculation to three significant digits. From this converged reference value we will define two regions. A first region

$$\mathcal{R}_\epsilon := \left[ (1 - \epsilon/100) \hat{D}_{31}^r, (1 + \epsilon/100) \hat{D}_{31}^r \right] \quad (58)$$

for calculations that deviate less than or equal to an  $\epsilon\%$  with respect to  $\hat{D}_{31}^r$ . This interval will be used for selecting the resolutions through the following convergence criteria. We say that, for fixed  $(N_\theta, N_\zeta, N_\xi)$  and  $\epsilon$ , a calculation  $\hat{D}_{31} \in \mathcal{R}_\epsilon$  is sufficiently converged if two conditions are satisfied

- (i) Spatially converged calculations with  $N'_\xi \geq N_\xi$  belong to  $\mathcal{R}_\epsilon$ .
- (ii) Increasing  $N_\theta$  and  $N_\zeta$  while keeping  $N_\xi$  constant produces calculations which belong to  $\mathcal{R}_\epsilon$ .

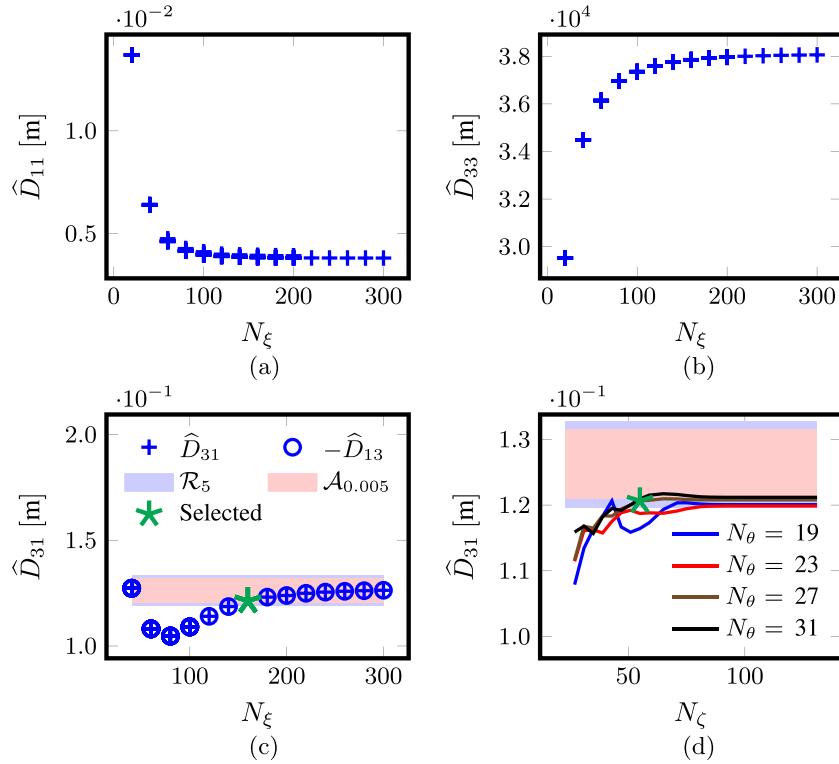
Condition (i) is used to select the number of Legendre modes  $N_\xi$  and condition (ii) is used to select the values of  $N_\theta$  and  $N_\zeta$  once  $N_\xi$  is fixed.

Additionally, we define a second interval

$$\mathcal{A}_\epsilon := \left[ \hat{D}_{31}^r - \epsilon, \hat{D}_{31}^r + \epsilon \right] \quad (59)$$

to distinguish which calculations are at a distance smaller than or equal to  $\epsilon$  from  $\hat{D}_{31}^r$ . The reason to have two different regions is that for stellarators close to QI, the relative convergence criteria can become too demanding (the smaller  $\hat{D}_{31}^r$  is, the narrower  $\mathcal{R}_\epsilon$  becomes). Nevertheless, for optimizing QI configurations, it is sufficient to ensure that  $|\hat{D}_{31}|$  is sufficiently small. If the absolute error is much smaller than a value of  $|\hat{D}_{31}|$  that can be considered sufficiently small, the calculation is converged for optimization purposes. We will use this interval for two purposes: first to give a visual idea of how narrow  $\mathcal{R}_\epsilon$  becomes. Second, to show that if  $\mathcal{R}_\epsilon$  is very small, it is easier to satisfy an absolute criteria than a relative one.

Figure 1 shows the convergence of monoenergetic coefficients with the number of Legendre modes for W7-X EIM when  $\hat{E}_r = 0$ . From figures 1(a) and (b) we see that the radial transport and parallel conductivity coefficients converge monotonically with  $N_\xi$ . On the other hand, the bootstrap current coefficient is more difficult to converge as can be seen on figure 1(c). As a sanity check, the fulfilment of the Onsager symmetry relation  $\hat{D}_{31} = -\hat{D}_{13}$  is included. The converged reference value  $\hat{D}_{31}^r$  is the spatially converged calculation for  $N_\xi = 380$ . Defining a region of relative convergence of  $\epsilon = 5\%$ , allows to select a resolution of  $N_\xi = 140$  Legendre modes to satisfy condition (i). The selection is indicated with a five-pointed green star. Note that for this case,



**Figure 2.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for W7X-EIM at the surface labelled by  $\psi/\psi_{\text{lcfs}} = 0.200$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r = 3 \cdot 10^{-4} \text{ kV} \cdot \text{s m}^{-2}$ .

an absolute deviation of 0.005 m from  $\hat{D}_{31}^r$  is slightly more demanding than the relative condition. This absolute deviation is selected as the 5% of  $\hat{D}_{31} \sim 0.1$  m, which can be considered a small value of  $\hat{D}_{31}$ . From figure 1(d) we choose the resolutions  $(N_\theta, N_\zeta) = (23, 55)$  to satisfy convergence condition (ii).

The case of W7-X EIM with  $\hat{E}_r \neq 0$  is shown in figure 2. We note from figure 2(c) that obtaining sufficiently converged results for the region  $\mathcal{R}_5$  is more difficult than in the case without radial electric field. For this case, the sizes of the intervals  $\mathcal{A}_{0.005}$  and  $\mathcal{R}_5$  are almost the same. This is in part due to the fact that the  $\hat{D}_{31}$  coefficient is smaller in absolute value and thus, the region  $\mathcal{R}_5$  is narrower. We select  $N_\xi = 160$  to satisfy condition (i). The selection  $(N_\theta, N_\zeta) = (27, 55)$  satisfies condition (ii) as shown in figure 2(d).

The convergence curves for the case of W7-X KJM when  $\hat{E}_r = 0$  are shown in figure 3. Due to the smallness of  $\hat{D}_{31}^r$ , the amplitude of the region  $\mathcal{R}_5$  is much narrower than in the EIM case, being of order  $10^{-3}$ . It is so narrow that the absolute value region  $\mathcal{A}_{0.005}$  contains the relative convergence region. It is shown in figure 3(c) that taking  $N_\xi = 140$  is sufficient to satisfy condition (i). According to the convergence curves plotted in figure 3(d), selecting  $(N_\theta, N_\zeta) = (23, 63)$  ensures satisfying condition (ii).

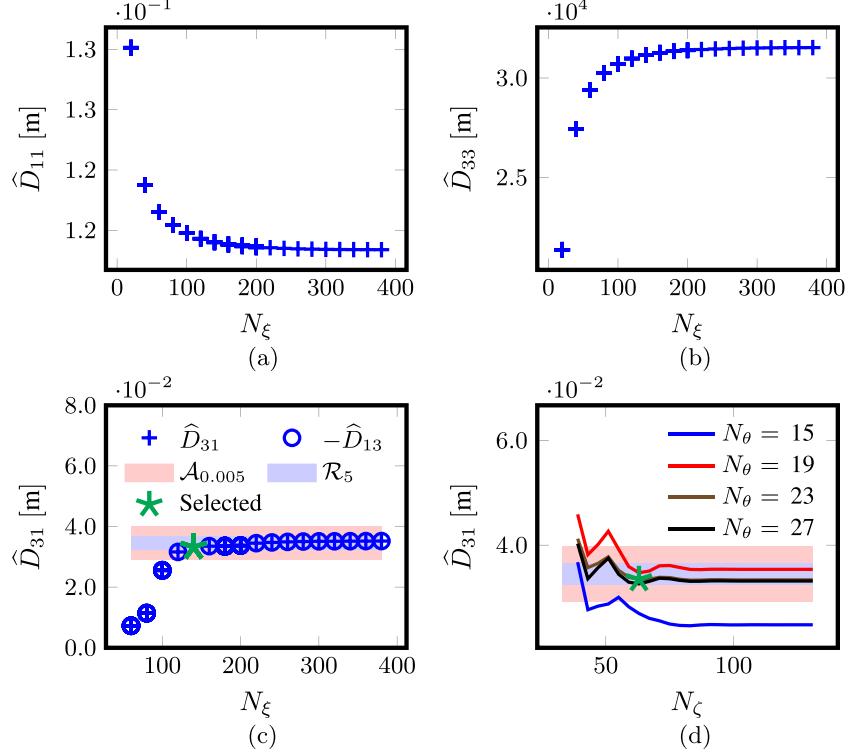
The case of W7-X KJM for finite  $\hat{E}_r$  is shown in figure 4. The selection of  $N_\xi = 180$  Legendre modes, indicated in figure 4(c), satisfies convergence condition (i). As shown in figure 4(d), condition (ii) is satisfied by the selection  $(N_\theta, N_\zeta) = (19, 79)$ .

The convergence of monoenergetic coefficients for CIEMAT-QI without  $\hat{E}_r$  is shown in figure 5. Note that as in the W7-X KJM case at this regime, the region of absolute error  $\mathcal{A}_{0.005}$  is bigger than the relative one. As the monoenergetic coefficients are smaller, we relax the relative convergence parameter to  $\epsilon = 7\%$ . In figure 5(c) we see that the region of 7% of deviation  $\mathcal{R}_7$  is quite narrow and that selecting  $N_\xi = 180$  satisfies condition (i). To satisfy condition (ii), we choose the resolutions  $(N_\theta, N_\zeta) = (15, 119)$  as shown in figure 5(d).

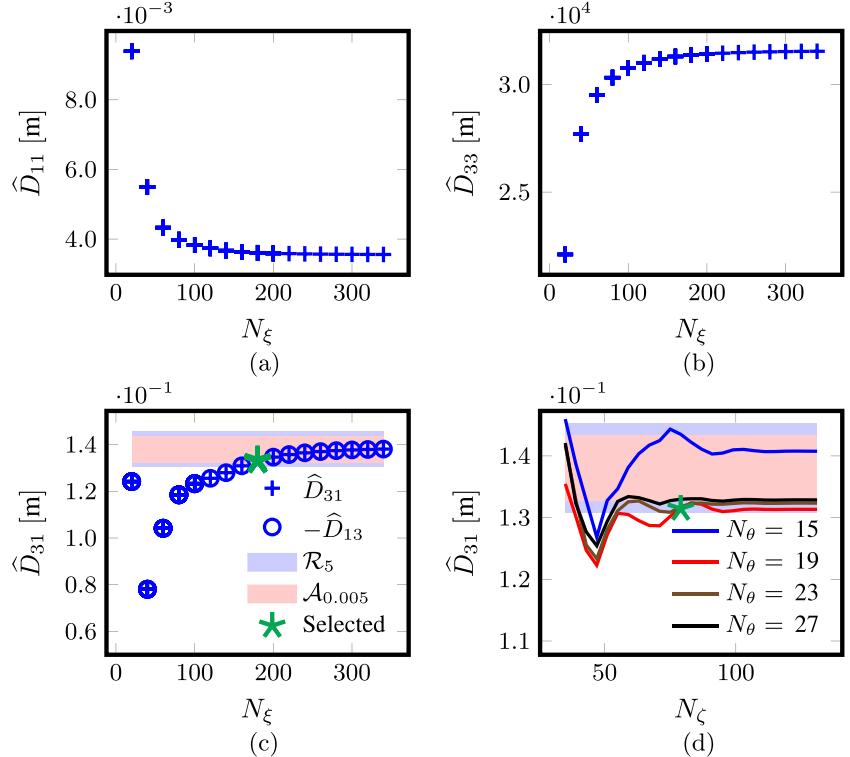
Finally, the case of CIEMAT-QI with  $\hat{E}_r \neq 0$  is shown in figure 6. Looking at figure 6(c) we can check that taking  $N_\xi = 180$  satisfies condition (i) for the region  $\mathcal{R}_7$  of 7% of deviation. In this case, the region of absolute error  $\mathcal{A}_{0.001}$  is five times smaller than in the rest of cases and is still bigger than the relative error region. As shown in figure 6(d), the selection  $(N_\theta, N_\zeta) = (15, 119)$  satisfies condition (ii).

#### 4.2. Code performance

In this subsection we will compare MONKES and DKES performance in terms of the wall-clock time and describe MONKES scaling properties. For the wall-clock time comparison, a convergence study (similar to the one explained in section 4.1) has been carried out with DKES on appendix E. This convergence study is done to compare the wall-clock times between MONKES and DKES for the same level of relative convergence with respect to  $\hat{D}_{31}^r$ . The comparison is displayed in table 2 along with the minimum number of Legendre modes for



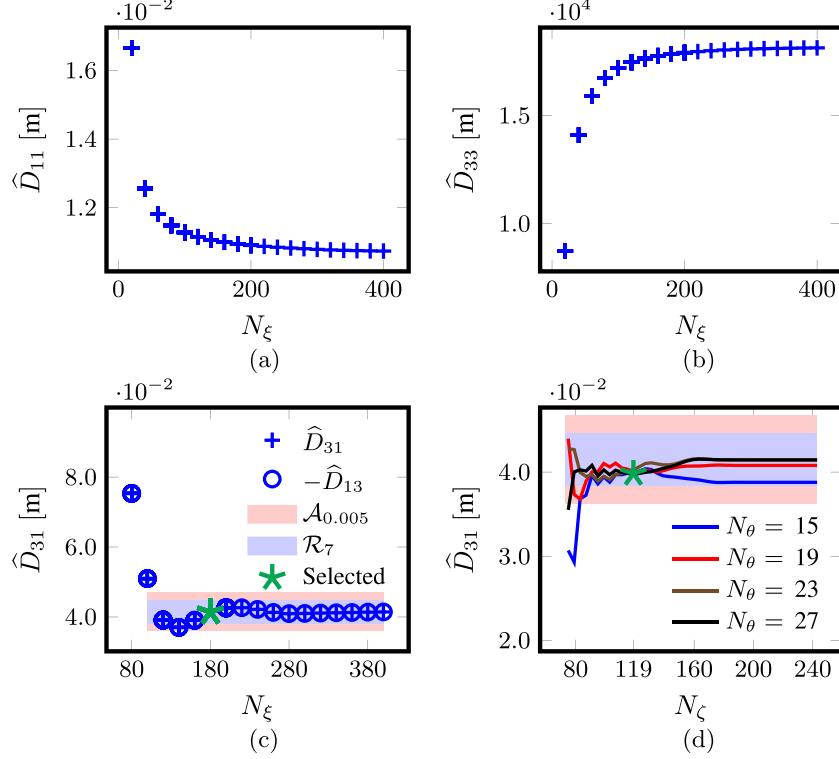
**Figure 3.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for W7X-KJM at the surface labelled by  $\psi/\psi_{\text{lfcs}} = 0.204$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ .



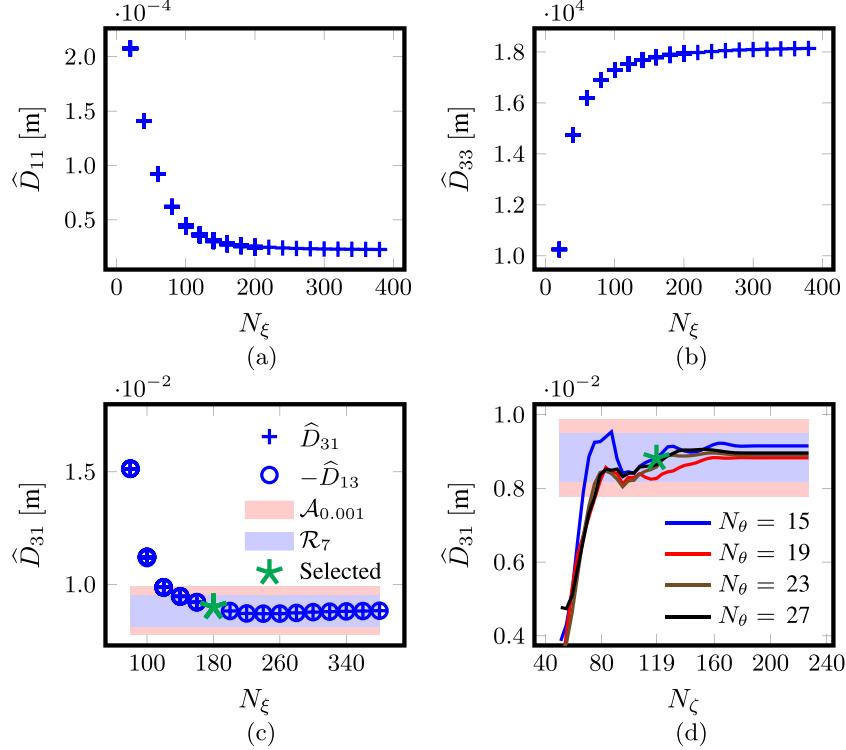
**Figure 4.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for W7X-KJM at the surface labelled by  $\psi/\psi_{\text{lfcs}} = 0.204$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 3 \cdot 10^{-4} \text{ kV} \cdot \text{s m}^{-2}$ .

which DKES results satisfy convergence condition (i). In all six cases, MONKES is faster than DKES despite using more Legendre modes. Even for W7-X EIM, in which we have

taken  $N_\xi = 40$  for DKES calculations with finite  $\hat{E}_r$ , MONKES is  $\sim 4$  times faster using almost four times the number of Legendre modes. For the W7-X EIM case without radial



**Figure 5.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for CIEMAT-QI at the surface labelled by  $\psi/\psi_{\text{lcfs}} = 0.25$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ .



**Figure 6.** Convergence of monoenergetic coefficients with the number of Legendre modes  $N_\xi$  for CIEMAT-QI at the surface labelled by  $\psi/\psi_{\text{lcfs}} = 0.25$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 10^{-3} \text{ kV} \cdot \text{s m}^{-2}$ .

electric field, the speed-up is also of 4. For the high mirror configuration, MONKES is  $\sim 20$  times faster than DKES without  $\hat{E}_r$  and  $\sim 10$  times faster than DKES when  $\hat{E}_r \neq 0$ . In the case

of CIEMAT-QI, MONKES is more than  $\sim 13$  times faster than DKES without radial electric field. In the case with finite  $\hat{E}_r$ , MONKES calculations are around 64 times faster than DKES

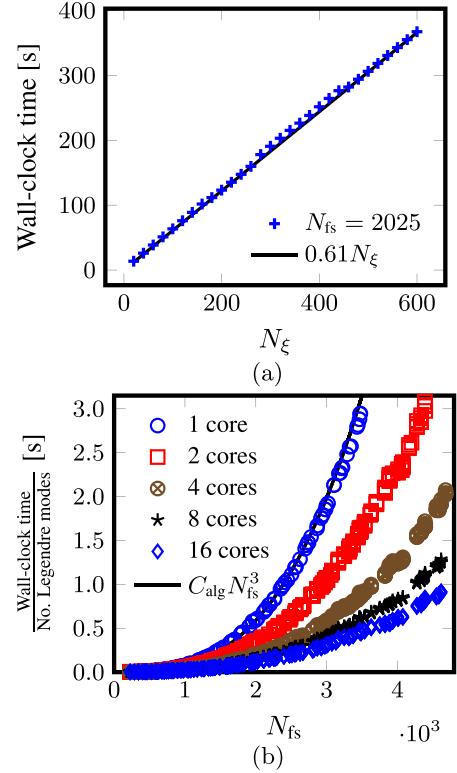
**Table 2.** Comparison between the wall-clock time of DKES and MONKES.

| Case                            | $N_\xi^{\text{DKES}}$ | $N_\xi^{\text{MONKES}}$ | $t_{\text{clock}}^{\text{DKES}}$ | $t_{\text{clock}}^{\text{MONKES}}$ |
|---------------------------------|-----------------------|-------------------------|----------------------------------|------------------------------------|
| W7X-EIM<br>$\hat{E}_r = 0$      | 80                    | 140                     | 90 s                             | 22 s                               |
| W7X-EIM<br>$\hat{E}_r \neq 0$   | 40                    | 160                     | 172 s                            | 35 s                               |
| W7X-KJM<br>$\hat{E}_r = 0$      | 160                   | 140                     | 698 s                            | 31 s                               |
| W7X-KJM<br>$\hat{E}_r \neq 0$   | 60                    | 180                     | 421 s                            | 47 s                               |
| CIEMAT-QI<br>$\hat{E}_r = 0$    | 160                   | 180                     | 1060 s                           | 76 s                               |
| CIEMAT-QI<br>$\hat{E}_r \neq 0$ | 160                   | 180                     | 4990 s                           | 76 s                               |

ones. One calculation of MONKES takes less than a minute and a half and the same calculation with DKES requires waiting for almost an hour and a half. The disparity of wall-clock times reflects the superiority at low collisionality of the block tridiagonal algorithm used by MONKES when compared to the iterative method used by DKES to solve the variational principle. The conjugate gradient method used by DKES converges slower (i.e. requires more iterations) when  $\hat{\nu}$  decreases while the performance of the block tridiagonal method does not depend on  $\hat{\nu}$ . We point out that the wall-clock times for all the calculations shown are those from one of the partitions of CIEMAT's cluster XULA. Specifically, partition number 2 has been used, whose nodes run with Intel Xeon Gold 6254 cores at 3.10 GHz.

We next check that the arithmetic complexity of the algorithm described in section 3 holds in practice. The scaling of MONKES with the number of Legendre modes  $N_\xi$  and the number of points in which the flux-surface is discretized is shown in figure 7. To demonstrate the linear scaling, the wall-clock time as a function of  $N_\xi$  for  $N_{\text{fs}} = 2025$  points is represented in figure 7(a) and compared with the line of slope 0.61 s per Legendre mode. As can be seen in figure 7(b), the wall-clock time (per Legendre mode) scales cubically with the number of points in which the flux-surface is discretized  $N_{\text{fs}}$ . As it was mentioned at the end of section 3, the constant  $C_{\text{alg}}$  in a single core is sufficiently small to give fast and accurate calculations up to  $\hat{\nu} \sim 10^{-5} \text{ m}^{-1}$ . We have plotted in figure 7(b) the cubic fit  $C_{\text{alg}}N_{\text{fs}}^3$ , where  $C_{\text{alg}} = 0.61(1/2025)^3 \sim 7 \cdot 10^{-11} \text{ s}$ .

As the LAPACK library is multithreaded and allows to parallelize the linear algebra operations through several cores, the scaling of MONKES when running in parallel is represented. Additionally, for the resolutions selected in section 4.1, we display in table 3 the wall-clock time when running MONKES using several cores in parallel. Note that for the W7-X cases, which require a smaller value of  $N_{\text{fs}}$ , the speed-up stalls at 8 cores. For CIEMAT-QI, that requires discretizing the flux-surface on a finer mesh, this does not happen in the range of cores considered.



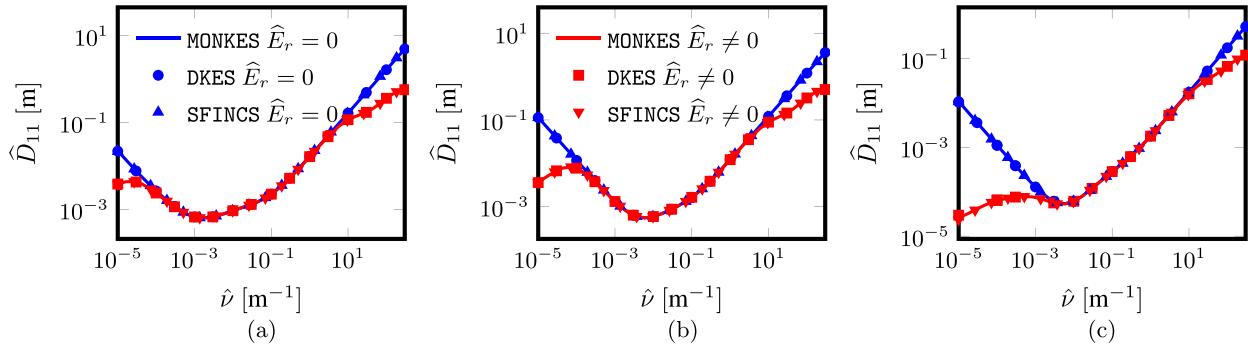
**Figure 7.** Scaling of MONKES wall-clock time. (a) Linear scaling with the number of Legendre modes for  $N_{\text{fs}} = 27 \times 75 = 2025$  discretization points. (b) Cubic scaling with  $N_{\text{fs}}$  for different number of cores used.

**Table 3.** Wall-clock time of MONKES in seconds for the selected triplets  $(N_\theta, N_\zeta, N_\xi)$  when running in several cores.

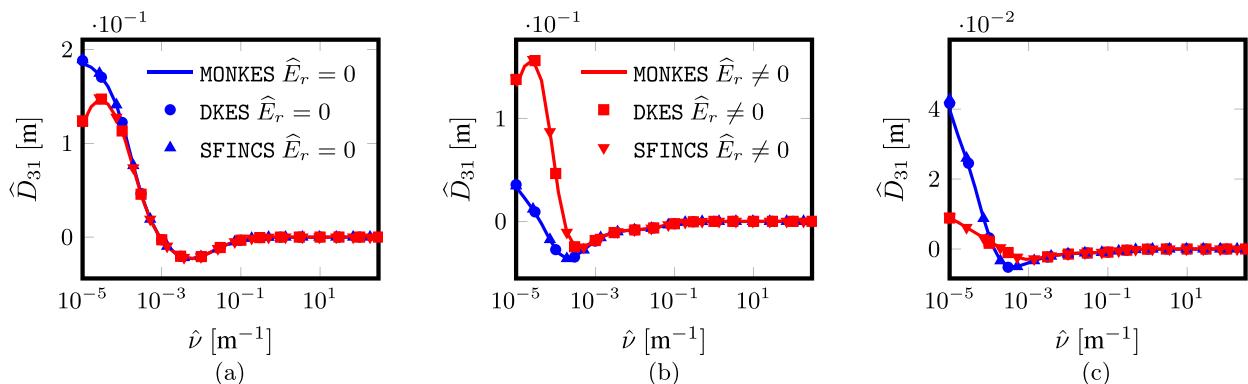
| Case \ No. cores             | 1  | 2  | 4  | 8  | 16 |
|------------------------------|----|----|----|----|----|
| W7X-EIM $\hat{E}_r = 0$      | 22 | 13 | 8  | 5  | 5  |
| W7X-EIM $\hat{E}_r \neq 0$   | 40 | 20 | 12 | 8  | 6  |
| W7X-KJM $\hat{E}_r = 0$      | 33 | 17 | 12 | 7  | 7  |
| W7X-KJM $\hat{E}_r \neq 0$   | 46 | 17 | 13 | 7  | 7  |
| CIEMAT-QI $\hat{E}_r = 0$    | 78 | 45 | 29 | 21 | 16 |
| CIEMAT-QI $\hat{E}_r \neq 0$ | 78 | 45 | 29 | 21 | 16 |

#### 4.3. Benchmark of the monoenergetic coefficients

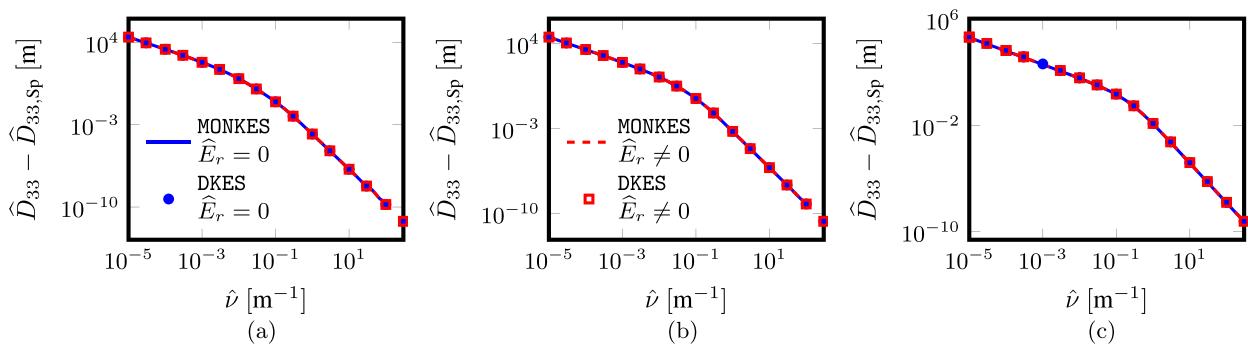
Once we have chosen the resolutions  $(N_\theta, N_\zeta, N_\xi)$  for each case, we need to verify that these selections indeed provide sufficiently accurate calculations of all the monoenergetic coefficients in the interval  $\hat{\nu} \in [10^{-5}, 300] \text{ m}^{-1}$ . It is instructive to recall what was mentioned at the beginning of section 4.1: that the number of Legendre modes required for converged calculations of the monoenergetic coefficients decreases when  $\hat{\nu}$  increases. Hence, the resolutions selected in section 4.1 also provide converged calculations for  $\hat{\nu} \geq 10^{-5}$ . For instance, for the W7X-EIM case and collisionality  $\hat{\nu} = 10^{-4}$ , taking  $N_\xi = 20$  is sufficient to have calculations converged up to 5% for zero and finite  $\hat{E}_r$ . In all cases, MONKES calculations of the  $\hat{D}_{11}$  and  $\hat{D}_{31}$  coefficients will



**Figure 8.** Calculation of  $\widehat{D}_{11}$  by MONKES, DKES and SFINCS for zero and finite  $\widehat{E}_r$ . (a) W7-X EIM at the surface  $\psi / \psi_{\text{lfcs}} = 0.200$ . (b) W7-X KJM at the surface  $\psi / \psi_{\text{lfcs}} = 0.204$ . (c) CIEMAT-QI at the surface  $\psi / \psi_{\text{lfcs}} = 0.250$ .



**Figure 9.** Calculation of  $\widehat{D}_{31}$  by MONKES, DKES and SFINCS for zero and finite  $\widehat{E}_r$ . (a) W7-X EIM at the surface  $\psi/\psi_{\text{lcfs}} = 0.200$ . (b) W7-X KJM at the surface  $\psi/\psi_{\text{lcfs}} = 0.204$ . (c) CIEMAT-QI at the surface  $\psi/\psi_{\text{lcfs}} = 0.250$ .



**Figure 10.** Calculation of  $\widehat{D}_{33}$  by MONKES and DKES for zero and finite  $\widehat{E}_r$ . (a) W7-X EIM at the surface  $\psi / \psi_{lcfs} = 0.200$ . (b) W7-X KJM at the surface  $\psi / \psi_{lcfs} = 0.204$ . (c) CIEMAT-QI at the surface  $\psi / \psi_{lcfs} = 0.250$ .

be benchmarked against converged calculations from DKES (see appendix E) and from SFINCS<sup>8</sup>. The parallel conductivity coefficient will be benchmarked only against DKES. The benchmarking of the coefficient  $\hat{D}_{11}$  for the six different cases is shown in figure 8. The result of the benchmark of the bootstrap current coefficient  $\hat{D}_{31}$  is shown in figure 9. Finally, the parallel conductivity coefficient  $\hat{D}_{33}$  is benchmarked in figure 10. Due to the weak effect of the radial electric field in the  $\hat{D}_{33}$  coefficient, the symbols for this plot have been

changed. In all cases, the agreement between MONKES, DKES and SFIMCS is almost perfect. Thus, we conclude that MONKES calculations of the monoenergetic coefficients are not only fast, but also accurate. Additionally, we can evaluate the level of optimization of the three configurations considered by inspecting these plots. In figures 8(a) and (b) is shown that the W7X-EIM configuration has smaller radial transport coefficient than the W7X-KJM configuration. Figures 9(a) and (b) show that the smaller radial transport of the W7X-EIM configuration comes at the expense of having larger bootstrap current coefficient. As shown in figures 8(c) and 9(c), the optimized stellarator CIEMAT-QI manages to achieve levels of radial transport similar or smaller than the W7X-EIM

<sup>8</sup> SFINCS calculations are converged up to 3% in the three independent variables.

configuration and a bootstrap current coefficient as low as the W7X-KJM configuration.

## 5. Conclusions and future work

In this paper we have presented the new code MONKES, which can provide fast and accurate calculations of the monoenergetic transport coefficients at low collisionality in a single core. By means of a thorough convergence study we have shown that it is possible to evaluate the monoenergetic coefficients in the  $1/\nu$  and  $\sqrt{\nu}\cdot\nu$  regimes in approximately 1 min. Besides, when there are sufficient computational resources available, the code can run even faster using several cores in parallel. A natural application is the inclusion of MONKES in a stellarator optimization suite. MONKES rapid calculations will allow direct optimization of the bootstrap current and radial transport from low collisionalities (typical of the  $1/\nu$  and  $\sqrt{\nu}\cdot\nu$ ) to moderate collisionalities (typical of the plateau regime). The low collisionality regimes are important in reactor relevant scenarios while the plateau regime can be important close to the edge, where the plasma is cooler. Massive evaluation of configurations to study the parametric dependence of  $\hat{D}_{31}$  or other coefficients on specific quantities of the magnetic configuration can also be done. Another application is its inclusion in predictive transport frameworks, which require neoclassical calculations to determine the evolution of plasma profiles. The neoclassical quantities required for these simulations can be calculated using MONKES.

Equation (1), solved by MONKES, includes a collision operator which does not preserve momentum. An important continuation of this work would be the implementation of momentum-correction techniques, such as the ones presented in [31–34]. As each calculation from MONKES can be executed in a single core, the scan in  $\nu$  (i.e. in  $\hat{\nu}$ ) required to perform the integrals of the monoenergetic coefficients is parallelizable. Therefore, it seems possible for the near future to obtain fast calculations of neoclassical transport with a model collision operator that preserves momentum. With this minor extension, MONKES could also be used for self-consistent optimization of magnetic fields in a similar manner to [16] for general geometry.

## Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No. 101052200—EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. This research was supported in part by Grant PID2021-123175NB-I00, Ministerio de Ciencia e Innovación, Spain. M L was supported by the U.S. Department of Energy, Office of Science,

Office of Fusion Energy Science, under Award Number DE-FG02-93ER54197. This work was supported by the U.S. Department of Energy under Contract Number DE-AC02-09CH11466. The United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The authors would like to thank C.D. Beidler for useful suggestions on the usage of DKES.

## Appendix A. Onsager symmetry

In this appendix we will prove that the monoenergetic coefficients  $\hat{D}_{ij}$  defined by (26) satisfy Onsager symmetry relations whenever there is no electric field  $E_\psi = 0$  or the magnetic field possesses stellarator symmetry. For this, we will prove a more general result involving linear equations defined in some domain (phase-space)  $\mathcal{S}$ . Suppose we have a space  $\mathcal{F}_\mathcal{S}$  of functions from  $\mathcal{S}$  to  $\mathbb{R}$  with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{S}}$  and a set of linear equations

$$\mathcal{V}f_j - \mathcal{C}f_j = s_j, \quad (\text{A.1})$$

for  $j = 1, 2, \dots, N_e$  where  $s_j \in \mathcal{F}_\mathcal{S}$  and the linear operators  $\mathcal{C}$  and  $\mathcal{V}$  are respectively symmetric and antisymmetric with respect to  $\langle \cdot, \cdot \rangle_{\mathcal{S}}$ . Namely,

$$\langle \mathcal{C}f, g \rangle_{\mathcal{S}} = \langle f, \mathcal{C}g \rangle_{\mathcal{S}}, \quad (\text{A.2})$$

$$\langle \mathcal{V}f, g \rangle_{\mathcal{S}} = -\langle f, \mathcal{V}g \rangle_{\mathcal{S}}. \quad (\text{A.3})$$

Now, we define the scalars

$$\mathcal{D}_{ij} := \langle s_i, f_j \rangle_{\mathcal{S}} \quad (\text{A.4})$$

for  $i, j = 1, 2, \dots, N_e$ .

Additionally, we define a property  $\mathcal{P}$  to be a map which associates to each  $f \in \mathcal{F}_\mathcal{S}$  a function  $\mathcal{P}f \in \mathcal{F}_\mathcal{S}$  and is idempotent<sup>9</sup>. Any function  $f \in \mathcal{F}_\mathcal{S}$  can be splitted in its even  $f^+$  and odd  $f^-$  portions with respect to the property  $\mathcal{P}$  as follows

$$f^\pm := \frac{1}{2}(f \pm \mathcal{P}f), \quad (\text{A.5})$$

satisfying  $\mathcal{P}f^\pm = \pm f^\pm$ . Without loss of generality, we assume that  $N^+ \leq N_e$  sources  $s_j$  in (A.1) are even with respect to  $\mathcal{P}$  and the remaining  $N^- := N_e - N^+$  sources are odd.

The coefficients  $\mathcal{D}_{ij}$  satisfy Onsager symmetry relations if three (sufficient) conditions are satisfied.

- (i) Even and odd functions are mutually orthogonal  
 $\langle f^\pm, g^\mp \rangle_{\mathcal{S}} = 0$ . This implies that

$$\langle f, g \rangle_{\mathcal{S}} = \langle f^+, g^+ \rangle_{\mathcal{S}} + \langle f^-, g^- \rangle_{\mathcal{S}}. \quad (\text{A.6})$$

<sup>9</sup> This means that, for all  $f \in \mathcal{F}_\mathcal{S}$ ,  $\mathcal{P}\mathcal{P}f = f$ .

- (ii) The operator  $\mathcal{C}$  is even with respect to property  $\mathcal{P}$ . Explicitly,

$$(\mathcal{C}f)^\pm = \mathcal{C}f^\pm. \quad (\text{A.7})$$

- (iii) The operator  $\mathcal{V}$  is odd with respect to property  $\mathcal{P}$ . Explicitly,

$$(\mathcal{V}f)^\pm = \mathcal{V}f^\mp. \quad (\text{A.8})$$

When conditions (A.6)–(A.8) are satisfied we have the following Onsager symmetry relations.

- For fixed  $i$  and  $j$ , if  $s_i$  and  $s_j$  are both even,  $\mathcal{D}_{ij} = \mathcal{D}_{ji}$ . The proof is as follows

$$\begin{aligned} \mathcal{D}_{ij} &= \left\langle s_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i^-, f_j^+ \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i^-, \mathcal{V}f_j^+ \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i^-, \mathcal{C}f_j^- \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i, \mathcal{C}f_j \right\rangle_{\mathcal{S}}. \end{aligned}$$

As in the last equality, due to (A.2), the roles of  $i$  and  $j$  are interchangeable, we have that  $\mathcal{D}_{ij} = \mathcal{D}_{ji}$ .

- For fixed  $i$  and  $j$ , if  $s_i$  and  $s_j$  are both odd,  $\mathcal{D}_{ij} = \mathcal{D}_{ji}$ . The proof is as follows

$$\begin{aligned} \mathcal{D}_{ij} &= \left\langle s_i^-, f_j^- \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i^+, f_j^- \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^-, f_j^- \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i^+, \mathcal{V}f_j^- \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^-, f_j^- \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i^+, \mathcal{C}f_j^+ \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^-, f_j^- \right\rangle_{\mathcal{S}} \\ &= - \left\langle f_i, \mathcal{C}f_j \right\rangle_{\mathcal{S}}. \end{aligned}$$

As in the last equality, due to (A.2), the roles of  $i$  and  $j$  are interchangeable, we have that  $\mathcal{D}_{ij} = \mathcal{D}_{ji}$ .

- For fixed  $i$  and  $j$ , if  $s_i$  is even and  $s_j$  is odd,  $\mathcal{D}_{ij} = -\mathcal{D}_{ji}$ . The proof is as follows

$$\begin{aligned} \mathcal{D}_{ij} &= \left\langle s_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i^-, f_j^+ \right\rangle_{\mathcal{S}} - \left\langle \mathcal{C}f_i^+, f_j^+ \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i^-, f_j^+ \right\rangle_{\mathcal{S}} - \left\langle f_i^+, \mathcal{C}f_j^+ \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i^-, f_j^+ \right\rangle_{\mathcal{S}} - \left\langle f_i^+, \mathcal{V}f_j^- \right\rangle_{\mathcal{S}} \\ &= \left\langle \mathcal{V}f_i, f_j \right\rangle_{\mathcal{S}}. \end{aligned}$$

As in the last equality, due to (A.3), interchanging the roles of  $i$  and  $j$  switches signs, we have that  $\mathcal{D}_{ij} = -\mathcal{D}_{ji}$ .

The equation (18) can be written in the form of (A.1) by setting the operators to be

$$\mathcal{V} = \xi \mathbf{b} \cdot \nabla + \nabla \cdot \mathbf{b} \frac{1-\xi^2}{2} \frac{\partial}{\partial \xi} - \frac{\widehat{E}_\psi}{\langle B^2 \rangle} \mathbf{B} \times \nabla \psi \cdot \nabla, \quad (\text{A.9})$$

$$\mathcal{C} = \hat{\nu} \mathcal{L}, \quad (\text{A.10})$$

and the inner product

$$\langle f, g \rangle_{\mathcal{S}} := \left\langle \int_{-1}^1 fg d\xi \right\rangle. \quad (\text{A.11})$$

With these definitions, properties (A.2) and (A.3)<sup>10</sup> are satisfied and  $\mathcal{D}_{ij} = \widehat{D}_{ij}$ . It is interesting to remark that the antisymmetry property (A.3) of  $\mathcal{V}$  implies that the diagonal monoenergetic coefficients  $\widehat{D}_{ii}$  are always positive. Note first that (A.3) implies  $\langle f, \mathcal{V}f \rangle_{\mathcal{S}} = 0$  for any  $f \in \mathcal{F}_{\mathcal{S}}$ . This implies that  $\widehat{D}_{ii} = -\langle f_i, \hat{\nu} \mathcal{L} f_i \rangle_{\mathcal{S}}$  and, as  $\mathcal{L}$  is a negative operator (its eigenvalues are all negative or zero, see appendix B),  $\widehat{D}_{ii} \geq 0$ . Also note that properties (A.2) and (A.3) imply that  $\langle \hat{\nu} \mathcal{L} f_j, 1 \rangle_{\mathcal{S}} = 0$  and  $\langle \mathcal{V}f_j, 1 \rangle_{\mathcal{S}} = 0$ . Thus, the image of the drift-kinetic equation (18) is constrained by  $\langle s_j, 1 \rangle_{\mathcal{S}} = 0$ .

Now we distinguish the two cases for which the monoenergetic coefficients  $\widehat{D}_{ij}$  satisfy Onsager symmetry relations. Apart from the velocity coordinate  $\xi$ , we will use Boozer coordinates  $(\theta, \zeta)$ .

- (i) If  $E_\psi = 0$ , the property is defined as

$$\mathcal{P}f(\theta, \zeta, \xi) = f(\theta, \zeta, -\xi). \quad (\text{A.12})$$

It is straightforward to check that for this property, conditions (A.6)–(A.8) are satisfied. Also,  $s_1 = s_1^+$ ,  $s_2 = s_2^+$  and  $s_3 = s_3^-$ . Hence, we have  $\widehat{D}_{12} = \widehat{D}_{21}$ ,  $\widehat{D}_{13} = -\widehat{D}_{31}$  and  $\widehat{D}_{23} = -\widehat{D}_{32}$ .

- (ii) When  $E_\psi$  is not necessarily zero, we define the property  $\mathcal{P}$  as the one that defines stellarator symmetry [40]

$$\mathcal{P}f(\theta, \zeta, \xi) = f(-\theta, -\zeta, \xi) \quad (\text{A.13})$$

and we have assumed without loss of generality that the planes of symmetry are  $\theta = 0$  and  $\zeta = 0$ . Thus, when the magnetic field is stellarator-symmetric  $B = B^+$ . In this case, using (29), (31) and (32) it is straightforward to check<sup>11</sup> that conditions (A.6)–(A.8) are satisfied. Besides,  $s_1 = s_1^-$ ,  $s_2 = s_2^-$  and  $s_3 = s_3^+$ . Hence, we have  $\widehat{D}_{12} = \widehat{D}_{21}$ ,  $\widehat{D}_{13} = -\widehat{D}_{31}$  and  $\widehat{D}_{23} = -\widehat{D}_{32}$ .

Note that for equation (18), the Onsager symmetry relation  $\widehat{D}_{12} = \widehat{D}_{21}$  is trivial as  $s_1 = s_2$ , which implies  $f_1 = f_2$  and thus  $\widehat{D}_{12} = \widehat{D}_{21} = \widehat{D}_{11} = \widehat{D}_{22}$ ,  $\widehat{D}_{31} = \widehat{D}_{32}$  and  $\widehat{D}_{13} = \widehat{D}_{23}$ .

<sup>10</sup> As  $\nabla \cdot \mathbf{v}_E = 0$ , the operator  $\mathcal{V}$  can be written in divergence form. For the symmetry of  $\mathcal{L}$  see appendix B.

<sup>11</sup> Note that derivatives along  $\theta$  and  $\zeta$  switch parities with respect to the stellarator symmetry property, i.e.  $\partial f^\pm / \partial \theta = (\partial f^\pm / \partial \theta)^\mp$  and  $\partial f^\pm / \partial \zeta = (\partial f^\pm / \partial \zeta)^\mp$ . Also, as for stellarator-symmetric fields,  $\sqrt{g} = \sqrt{g}^+$  the flux-surface average satisfies  $\langle f^- \rangle = 0$ .

Nevertheless, if the definition of  $s_1$  and  $s_2$  was different, as long as their parity is the same, the relation  $\widehat{D}_{12} = \widehat{D}_{21}$  would still hold.

## Appendix B. Legendre modes of the drift-kinetic equation

Legendre polynomials are the eigenfunctions of the Sturm–Liouville problem in the interval  $\xi \in [-1, 1]$  defined by the differential equation

$$2\mathcal{L}P_k(\xi) = -k(k+1)P_k(\xi), \quad (\text{B.1})$$

and regularity boundary conditions at  $\xi = \pm 1$

$$(1 - \xi^2) \frac{dP_k}{d\xi} \Big|_{\xi=\pm 1} = 0, \quad (\text{B.2})$$

where  $k \geq 0$  is an integer.

As  $\mathcal{L}$  has a discrete spectrum and is self-adjoint with respect to the inner product

$$\langle f, g \rangle_{\mathcal{L}} := \int_{-1}^1 fg d\xi, \quad (\text{B.3})$$

in the space of functions that satisfy the regularity condition,  $\{P_k\}_{k=0}^{\infty}$  is an orthogonal basis satisfying  $\langle P_j, P_k \rangle_{\mathcal{L}} = 2\delta_{jk}/(2k+1)$ . Hence, these polynomials satisfy the three-term recurrence formula

$$(2k+1)\xi P_k(\xi) = (k+1)P_{k+1}(\xi) + kP_{k-1}(\xi), \quad (\text{B.4})$$

obtained by Gram–Schmidt orthogonalization. Starting from the initial values  $P_0 = 1$  and  $P_1 = \xi$ , the recurrence defines the rest of the Legendre polynomials. Additionally, they satisfy the differential identity

$$(1 - \xi^2) \frac{dP_k}{d\xi} = kP_{k-1}(\xi) - k\xi P_k(\xi). \quad (\text{B.5})$$

Identities (B.4) and (B.5) are useful to represent tridiagonally the left-hand side of equation (18) when we use the expansion (33). The  $k$ -th Legendre mode of the term  $\xi \mathbf{b} \cdot \nabla f$  is expressed in terms of the modes  $f^{(k-1)}$  and  $f^{(k+1)}$  using (B.4)

$$\begin{aligned} \langle \xi \mathbf{b} \cdot \nabla f, P_k \rangle_{\mathcal{L}} &= \frac{2}{2k+1} \left[ \frac{k}{2k-1} \mathbf{b} \cdot \nabla f^{(k-1)} \right. \\ &\quad \left. + \frac{k+1}{2k+3} \mathbf{b} \cdot \nabla f^{(k+1)} \right]. \end{aligned} \quad (\text{B.6})$$

Combining both (B.4) and (B.5) allows to express the  $k$ -th Legendre mode of the mirror term  $\nabla \cdot \mathbf{b}((1 - \xi^2)/2) \partial f / \partial \xi$  in terms of the modes  $f^{(k-1)}$  and  $f^{(k+1)}$  as

$$\begin{aligned} &\left\langle \frac{1}{2}(1 - \xi^2) \nabla \cdot \mathbf{b} \frac{\partial f}{\partial \xi}, P_k \right\rangle_{\mathcal{L}} \\ &= \frac{\mathbf{b} \cdot \nabla \ln B}{2k+1} \left[ \frac{k(k-1)}{2k-1} f^{(k-1)} - \frac{(k+1)(k+2)}{2k+3} f^{(k+1)} \right], \end{aligned} \quad (\text{B.7})$$

where we have also used  $\nabla \cdot \mathbf{b} = -\mathbf{b} \cdot \nabla \ln B$ . The term proportional to  $\widehat{E}_{\psi}$  is diagonal in a Legendre representation

$$\begin{aligned} &\left\langle \frac{\widehat{E}_{\psi}}{\langle B^2 \rangle} \mathbf{B} \times \nabla \psi \cdot \nabla f, P_k \right\rangle_{\mathcal{L}} \\ &= \frac{2}{2k+1} \frac{\widehat{E}_{\psi}}{\langle B^2 \rangle} \mathbf{B} \times \nabla \psi \cdot \nabla f^{(k)}. \end{aligned} \quad (\text{B.8})$$

For the collision operator used in equation (18), as Legendre polynomials are eigenfunctions of the pitch-angle scattering operator, using (B.1) we obtain the diagonal representation

$$\langle \hat{\nu} \mathcal{L} f, P_k \rangle_{\mathcal{L}} = -\hat{\nu} \frac{k(k+1)}{2k+1} f^{(k)}. \quad (\text{B.9})$$

Finally, we briefly comment on why the truncation error from (33) implies that the solution to (34) and (38) is an approximation of the Legendre spectrum of the exact solution to (18) satisfying (13). For this, we will assume that the solution to (18) and (13) is unique (which it is, see appendix C). We denote this exact solution by  $f_{\text{ex}}$  and its Legendre modes by  $f_{\text{ex}}^{(k)}$ . The Legendre modes  $f_{\text{ex}}^{(k)}$  satisfy (34) for all values of  $k$ , including  $k > N_{\xi}$  and, in general,  $f_{\text{ex}}^{(N_{\xi}+1)} \neq 0$ . Denoting the error of the solution  $f^{(k)}$  to (34) and (38) by

$$E^{(k)} := f_{\text{ex}}^{(k)} - f^{(k)}, \quad (\text{B.10})$$

is easy to prove that

$$L_k E^{(k-1)} + D_k E^{(k)} + U_k E^{(k+1)} = 0, \quad (\text{B.11})$$

for  $k = 0, 1, \dots, N_{\xi} - 1$  and

$$L_{N_{\xi}} E^{(N_{\xi}-1)} + D_{N_{\xi}} E^{(N_{\xi})} = -U_{N_{\xi}} f_{\text{ex}}^{(N_{\xi}+1)}. \quad (\text{B.12})$$

Note that the system of equations constituted by (B.11) and (B.12) for the error is identical to (34) substituting  $f^{(k)}$  by  $E^{(k)}$  and  $s^{(k)}$  by  $-U_{N_{\xi}} f_{\text{ex}}^{(N_{\xi}+1)}$ . Hence, by assumption, the solution to (B.11) and (B.12) satisfying (38) is unique, implying that  $E^{(k)} \neq 0$  unless  $U_{N_{\xi}} f_{\text{ex}}^{(N_{\xi}+1)} = 0$ .

## Appendix C. Invertibility of the spatial differential operators

In this appendix we will study the invertibility of the left-hand-side of (34). We are only concerned in elucidating under which conditions the algorithm given in section 3 can be applied to solve (34). For instance, we will consider the possibility of the flux-surface being rational despite of the fact that (among other things) it may be inconsistent with the assumption that thermodynamical forces are a flux-function. We will conclude that the solution to (34) submitted to (38) is unique in ergodic flux-surfaces and also on rational flux-surfaces with  $E_{\psi} \neq 0$  and can be obtained with the aforementioned algorithm. In order to do this, we view  $L_k$ ,  $D_k$  and  $U_k$  as operators that act on  $\mathcal{F}$ ,

where  $\mathcal{F}$  is the space of smooth functions on the flux-surface equipped with the inner product

$$\langle f, g \rangle_{\mathcal{F}} = \frac{N_p}{4\pi^2} \oint \oint f \bar{g} d\theta d\zeta, \quad (\text{C.1})$$

where  $\bar{z}$  denotes the complex conjugate of  $z$  and the inner product induces a norm

$$\|f\|_{\mathcal{F}} := \sqrt{\langle f, f \rangle_{\mathcal{F}}}. \quad (\text{C.2})$$

In this setting  $L_k$ ,  $D_k$  and  $U_k$  are operators from  $\mathcal{F}$  to  $\mathcal{F}$  as all of their coefficients are smooth on the flux-surface. However, the operators  $L_k$  and  $U_k$  given by (35) and (37) do not have a uniquely defined inverse. This is a consequence of the fact that the parallel streaming operator  $\xi \mathbf{B} \cdot \nabla + \nabla \cdot \mathbf{B}(1 - \xi^2)/2 \partial/\partial\xi$  has a non trivial kernel comprised of functions  $g((1 - \xi^2)/B)$ . On the other hand, the operator  $D_k$  has a unique inverse for  $k \geq 1$ . For  $k = 0$ , the operator  $D_0$  is not invertible as it has a kernel comprised of functions  $g(B_\theta \theta + B_\zeta \zeta)$ .

Whether  $L_k$  and  $U_k$  are or not invertible can be determined studying the uniqueness of continuous solutions (on the flux-surface) to

$$\mathbf{B} \cdot \nabla f + \omega_k f = sB, \quad (\text{C.3})$$

for some  $s, \omega_k \in \mathcal{F}$ . Note that equations  $L_k f = ks/(2k - 1)$  and  $U_k f = (k + 1)s/(2k + 3)$  can be written in the form of equation (C.3) setting, respectively,  $\omega_k = (k - 1)\mathbf{B} \cdot \nabla \ln B/2$  and  $\omega_k = -(k + 2)\mathbf{B} \cdot \nabla \ln B/2$ . We will determine a condition for  $\omega_k$  which, if satisfied, equation (C.3) has a unique solution  $f \in \mathcal{F}$ .

The solution to equation (C.3) can be written as

$$f = (f_0 + K)\Phi, \quad (\text{C.4})$$

where

$$\mathbf{B} \cdot \nabla f_0 = 0, \quad (\text{C.5})$$

$$\mathbf{B} \cdot \nabla \Phi + \omega_k \Phi = 0, \quad (\text{C.6})$$

$$\mathbf{B} \cdot \nabla K = sB/\Phi. \quad (\text{C.7})$$

Equations (C.6) and (C.7) are integrated (along a field line) imposing  $\Phi|_p = 1$  and  $K|_p = 0$  at a point  $p$  of the field line. Note that  $f_0 = f|_p$  is an integration constant. Depending on the form of  $\omega_k$ ,  $f_0$  can or cannot be determined imposing continuity on the flux-surface. The solution to equation (C.6) can be written as

$$\Phi = \exp(-W_k), \quad (\text{C.8})$$

where  $\mathbf{B} \cdot \nabla W_k = \omega_k$  and is integrated imposing  $W_k|_p = 0$ . Note that this implies that  $\Phi \neq 0$  and that

$$-\mathbf{B} \cdot \nabla \left( \frac{1}{\Phi} \right) + \omega_k \frac{1}{\Phi} = 0. \quad (\text{C.9})$$

When  $\Phi \in \mathcal{F}$ , the left-hand side of (C.3) has a non trivial kernel (as an operator from  $\mathcal{F}$  to  $\mathcal{F}$ ). In order to proceed further, we employ coordinates  $(\alpha, l)$  where  $\alpha := \theta - \iota\zeta$  is a poloidal

angle that labels field lines and  $l$  is the length along magnetic field lines. Depending on the type of flux-surface there are two possible situations

- (i) For ergodic flux-surfaces,  $\iota \in \mathbb{R} \setminus \mathbb{Q}$  and satisfying (C.5) implies that  $f_0$  is a flux-function. The solution  $f$  to (C.3) is a differentiable function on the torus if  $\langle \mathbf{B} \cdot \nabla f \rangle = 0$ . Applying (equation (C.3)) combined with splitting (C.4) yields

$$\begin{aligned} f_0 \langle \omega_k \Phi \rangle &= \langle Bs \rangle - \langle K \omega_k \Phi \rangle \\ &= \langle \mathbf{B} \cdot \nabla (K \Phi) \rangle. \end{aligned} \quad (\text{C.10})$$

Hence, if  $\langle \omega_k \Phi \rangle \neq 0$ , equation (C.10) fixes the value of  $f_0$  so that  $f$  is continuous on the torus. Note that if  $\langle \omega_k \Phi \rangle \neq 0$ , by virtue of (C.6),  $\Phi$  is not univaluated and does not belong to  $\mathcal{F}$ . On the contrary, if  $f_0$  is free, then  $\Phi$  is a continuous function on the torus. Then, (C.10) implies that  $K\Phi$  is continuous on the torus when  $\Phi$  is. The function  $K$  is also continuous as long as  $sB$  belongs to the image of  $\mathbf{B} \cdot \nabla + \omega_k$ . Note that using (C.9) we can derive from (equation (C.3)/ $\Phi$ ) the solvability condition  $\langle sB/\Phi \rangle = 0$ .

- (ii) For rational flux-surfaces,  $\iota \in \mathbb{Q}$  and satisfying (C.5) implies that  $f_0(\alpha)$  depends on the field line chosen. At these surfaces, the field line labelled by  $\alpha$  closes on itself after a length  $L_c(\alpha)$ . If the solution  $f$  is continuous on the flux-surface, then  $\int_0^{L_c} \mathbf{B} \cdot \nabla f dl/B = 0$  for each field line. Applying  $\int_0^{L_c}$  equation (C.3)  $dl/B$  combined with splitting (C.4) yields

$$\begin{aligned} f_0(\alpha) \int_0^{L_c} \omega_k \Phi \frac{dl}{B} &= \int_0^{L_c} s dl - \int_0^{L_c} \omega_k K \Phi \frac{dl}{B} \\ &= \int_0^{L_c} \mathbf{B} \cdot \nabla (K \Phi) \frac{dl}{B}. \end{aligned} \quad (\text{C.11})$$

If  $\int_0^{L_c} \omega_k \Phi dl/B \neq 0$ , condition (C.11) fixes a unique value of  $f_0(\alpha)$  (for each field line) for which  $f$  is continuous on the torus. As for ergodic surfaces, if (C.11) does not fix  $f_0$ , then  $\Phi$  and  $K\Phi$  are continuous along field lines. Again,  $K$  is also continuous as long as  $sB$  belongs to the image of  $\mathbf{B} \cdot \nabla + \omega_k$ . Using (C.9) we can derive from  $\int_0^{L_c}$  equation (C.3)/ $\Phi dl/B$  the solvability condition  $\int_0^{L_c} sB/\Phi dl/B = 0$ .

Thus, we have seen that when  $\langle \omega_k \Phi \rangle = 0$  or  $\int_0^{L_c} \omega_k \Phi dl/B = 0$ , the operator  $\mathbf{B} \cdot \nabla + \omega_k$  from  $\mathcal{F}$  to itself is not one-to-one (it has a non trivial kernel comprised of multiples of  $\Phi$ ). Moreover, we have the solvability conditions  $\langle sB/\Phi \rangle = 0$  for ergodic surfaces and  $\int_0^{L_c} sB/\Phi dl/B = 0$  for rational surfaces. The existence of a solvability condition implies that  $\mathbf{B} \cdot \nabla + \omega_k$  is not onto. We can derive a simpler and equivalent condition for  $\omega_k$  from (C.8). Note that  $\Phi$  is continuous on the torus only when  $W_k$  is. As  $\mathbf{B} \cdot \nabla W_k = \omega_k$ , continuity of  $W_k$  along field lines imposes  $\langle \omega_k \rangle = 0$  on ergodic flux-surfaces and  $\int_0^{L_c} \omega_k dl/B = 0$  on rational ones. Hence, the operator  $\mathbf{B} \cdot \nabla + \omega_k$  is invertible if  $\langle \omega_k \rangle \neq 0$  or  $\int_0^{L_c} \omega_k dl/B \neq 0$ .

This result can be applied to determine that  $L_k$  and  $U_k$  are not invertible. For both  $L_k$  and  $U_k$ ,  $\omega_k \propto \mathbf{B} \cdot \nabla \ln B^\gamma$  for some

rational exponent  $\gamma$ . As  $B$  is continuous on the flux-surface we have for  $L_k$  and  $U_k$  that  $\int_0^{L_\alpha} \omega_k dl/B = 0$  or  $\langle \omega_k \rangle = 0$ , which means that neither  $L_k$  nor  $U_k$  are invertible.

Now we turn our attention to the invertibility of  $D_k$  for  $k \geq 1$ . For  $\widehat{E}_\psi = 0$ ,  $D_k$  is just a multiplicative operator and is clearly invertible when  $\widehat{\nu}_k, k \neq 0$ . For  $\widehat{E}_\psi \neq 0$ , the invertibility of  $D_k$  can be proven by studying the uniqueness of solutions to

$$\mathbf{B} \times \nabla \psi \cdot \nabla g - \widehat{\nu}_k g = -\frac{\langle B^2 \rangle}{\widehat{E}_\psi} s, \quad (\text{C.12})$$

where  $\widehat{\nu}_k = \widehat{\nu} k(k+1) \langle B^2 \rangle / 2 \widehat{E}_\psi$ . The procedure is very similar to the one carried out for  $L_k$  and  $U_k$ . First, we write the solution to equation (C.12) as

$$g = (g_0 + I) \Psi, \quad (\text{C.13})$$

where

$$\mathbf{B} \times \nabla \psi \cdot \nabla g_0 = 0, \quad (\text{C.14})$$

$$\mathbf{B} \times \nabla \psi \cdot \nabla \Psi - \widehat{\nu}_k \Psi = 0, \quad (\text{C.15})$$

$$\mathbf{B} \times \nabla \psi \cdot \nabla I = -\frac{\langle B^2 \rangle}{\widehat{E}_\psi} \frac{s}{\Psi}. \quad (\text{C.16})$$

Equations (C.15) and (C.16) are integrated along a integral curve of  $\mathbf{B} \times \nabla \psi$  imposing  $\Psi|_p = 1$  and  $I|_p = 0$  at the initial point  $p$  of integration. The integral curves of  $\mathbf{B} \times \nabla \psi$  are, in Boozer coordinates, straight lines  $B_\theta \theta + B_\zeta \zeta = \text{constant}$ . In order to proceed further, we change from Boozer angles  $(\theta, \zeta)$  to a different set of magnetic coordinates  $(\alpha, \varphi)$  using the linear transformation

$$\begin{bmatrix} \theta \\ \zeta \end{bmatrix} = \begin{bmatrix} (1 + \iota \delta)^{-1} & \iota \\ -\delta(1 + \iota \delta)^{-1} & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \varphi \end{bmatrix} \quad (\text{C.17})$$

where  $\delta = B_\theta/B_\zeta$ . In these coordinates  $\mathbf{B} = \nabla \psi \times \nabla \alpha$ ,  $B_\alpha = 0$  and

$$\mathbf{B} \times \nabla \psi \cdot \nabla = B^2 \frac{\partial}{\partial \alpha}. \quad (\text{C.18})$$

Depending on the rationality or irrationality of  $\delta$  we can distinguish two options

(i) If  $\delta \in \mathbb{R} \setminus \mathbb{Q}$ , satisfying (C.14) implies that  $g_0$  is a flux-function (the integral curves trace out the whole flux-surface). Note that if  $g$  is a differentiable function on the torus  $\langle \mathbf{B} \times \nabla \psi \cdot \nabla g \rangle = \langle \nabla \times (g \mathbf{B}) \cdot \nabla \psi \rangle = 0$ , where we have used  $\nabla \times \mathbf{B} \cdot \nabla \psi = 0$ . Taking (equation (C.12)) assuming that  $g$  is continuous on the flux-surface, combined with (C.13) gives

$$\begin{aligned} \langle \Psi \rangle g_0 &= \frac{\langle B^2 \rangle}{\widehat{\nu}_k \widehat{E}_\psi} \langle s \rangle - \langle I \Psi \rangle \\ &= \frac{1}{\widehat{\nu}_k} \langle \mathbf{B} \times \nabla \psi \cdot \nabla (I \Psi) \rangle. \end{aligned} \quad (\text{C.19})$$

Hence, if  $\langle \Psi \rangle \neq 0$ , continuity of  $g$  on the torus fixes the integration constant  $g_0$ .

(ii) If  $\delta \in \mathbb{Q}$ , satisfying (C.14) implies that  $g_0(\varphi)$  is a function of  $\varphi$ . Now the integral curves  $\varphi = \text{constant}$  close on itself after moving in  $\alpha$  an arc-length  $L_\alpha$ . In this scenario, if  $g$  is a differentiable function on the torus  $\int_0^{L_\alpha} \mathbf{B} \times \nabla \psi \cdot \nabla g d\alpha/B^2 = 0$ , where we have used (C.18). Thus, taking  $\int_0^{L_\alpha}$  equation (C.12)  $d\alpha/B^2$ , combined with (C.13) gives

$$\begin{aligned} g_0(\varphi) \int_0^{L_\alpha} \Psi \frac{d\alpha}{B^2} &= \frac{\langle B^2 \rangle}{\widehat{\nu}_k \widehat{E}_\psi} \int_0^{L_\alpha} s \frac{d\alpha}{B^2} - \int_0^{L_\alpha} I \Psi \frac{d\alpha}{B^2} \\ &= \frac{1}{\widehat{\nu}_k} \int_0^{L_\alpha} \mathbf{B} \times \nabla \psi \cdot \nabla (I \Psi) \frac{d\alpha}{B^2}. \end{aligned} \quad (\text{C.20})$$

Thus, if  $\int_0^{L_\alpha} \Psi d\alpha/B^2 \neq 0$  condition (C.20) fixes the value of  $g_0(\varphi)$  so that  $g$  is continuous on the flux-surface.

Similarly to what happened to  $\Phi$  when studying the invertibility of  $L_k$  and  $U_k$ , continuity of the solution implies that  $\Psi$  cannot be univaluated. We can write  $\Psi$  as

$$\Psi = \exp(-A_k), \quad (\text{C.21})$$

where  $\mathbf{B} \times \nabla \psi \cdot \nabla A_k = \widehat{\nu}_k$  and is integrated along with condition  $A_k|_p = 0$ . Using (C.18), we can write

$$A_k(\alpha, \varphi) = \widehat{\nu}_k \int_0^\alpha \frac{d\alpha'}{B^2(\alpha', \varphi)}. \quad (\text{C.22})$$

Note that  $A_k$  is monotonically crescent with  $\alpha$ , which means that  $\Psi$  cannot be univaluated. Besides, (C.21) implies  $\Psi > 0$ , which means that  $\langle \Psi \rangle \neq 0$  and  $\int_0^{L_\alpha} \Psi d\alpha/B^2 \neq 0$ . Thus, there is a unique value of the constant  $g_0$  which compensates the jumps in  $\Psi$  and  $I \Psi$  so that  $g = g_0 \Psi + I \Psi$  is continuous on the flux-surface. Hence,  $D_k$  is an invertible operator from  $\mathcal{F}$  to itself.

The inverse of  $D_k$  for  $k \geq 1$  and  $\widehat{E}_\psi \neq 0$  is defined by

$$D_k^{-1} s := (\mathcal{G}_0[s] + \mathcal{I}[s]) \Psi, \quad (\text{C.23})$$

where  $\mathcal{G}_0[s]$  and  $\mathcal{I}[s]$  denote the linear operators which define, respectively, the constant of integration and the solution to (C.16) with  $I|_p = 0$  for a given source term. Specifically,

$$\mathcal{I}[s](\alpha, \varphi) := -\frac{\langle B^2 \rangle}{\widehat{E}_\psi} \int_0^\alpha \frac{s(\alpha', \varphi)}{\Psi(\alpha', \varphi)} \frac{d\alpha'}{B^2(\alpha', \varphi)}, \quad (\text{C.24})$$

and

$$\mathcal{G}_0[s](\varphi) := \begin{cases} \text{If } \delta \in \mathbb{R} \setminus \mathbb{Q}: \\ \frac{2}{\widehat{\nu}_k(k+1)} \frac{\langle s \rangle}{\langle \Psi \rangle} - \frac{\langle \mathcal{I}[s] \Psi \rangle}{\langle \Psi \rangle}, \\ \text{If } \delta \in \mathbb{Q}: \\ \frac{2}{\widehat{\nu}_k(k+1)} \frac{\int_0^{L_\alpha} s \frac{d\alpha}{B^2}}{\int_0^{L_\alpha} \Psi \frac{d\alpha}{B^2}} - \frac{\int_0^{L_\alpha} \mathcal{I}[s] \Psi \frac{d\alpha}{B^2}}{\int_0^{L_\alpha} \Psi \frac{d\alpha}{B^2}}. \end{cases} \quad (\text{C.25})$$

Finally, we will study the invertibility of the operator  $\Delta_k$

$$\Delta_k = D_k - U_k \Delta_{k+1}^{-1} L_{k+1} \quad (\text{C.26})$$

assuming that  $\Delta_{k+1}$  is an invertible operator from  $\mathcal{F}$  to  $\mathcal{F}$ . For this, first, we note that in the space of functions of interest (smooth periodic functions on the torus), using a Fourier basis  $\{e^{i(m\theta+nN_p\zeta)}\}_{m,n\in\mathbb{Z}}$ , we can approximate any function  $f(\theta,\zeta) = \sum_{m,n\in\mathbb{Z}} \hat{f}_{mn} e^{i(m\theta+nN_p\zeta)} \in \mathcal{F}$  using an approximant  $\tilde{f}(\theta,\zeta)$

$$\tilde{f}(\theta,\zeta) = \sum_{-N \leq m,n \leq N} \hat{f}_{mn} e^{i(m\theta+nN_p\zeta)} \quad (\text{C.27})$$

truncating the modes with mode number greater than some positive integer  $N$  where

$$\hat{f}_{mn} = \left\langle f, e^{i(m\theta+nN_p\zeta)} \right\rangle_{\mathcal{F}} \left\| e^{i(m\theta+nN_p\zeta)} \right\|_{\mathcal{F}}^{-2} \quad (\text{C.28})$$

are the Fourier modes of  $f$ . Thus, we approximate  $\mathcal{F}$  using a finite dimensional subspace  $\mathcal{F}^N \subset \mathcal{F}$  consisting on all the functions of the form given by equation (C.27).

Hence, we can approximate  $D_k$ ,  $U_k$ ,  $\Delta_{k+1}$  and  $L_{k+1}$  restricted to  $\mathcal{F}^N$  (and therefore  $\Delta_k$ ) in equation (C.26) by operators  $D_k^N$ ,  $U_k^N$ ,  $\Delta_{k+1}^N$  and  $L_{k+1}^N$  that map any  $\tilde{f} \in \mathcal{F}^N$  to the projections of  $D_k \tilde{f}$ ,  $U_k \tilde{f}$ ,  $\Delta_{k+1} \tilde{f}$  and  $L_{k+1} \tilde{f}$  onto  $\mathcal{F}^N$ . The operators  $D_k^N$ ,  $U_k^N$ ,  $\Delta_{k+1}^N$  and  $L_{k+1}^N$  can be exactly represented (in a Fourier basis) by square matrices of size  $\dim \mathcal{F}^N$ . When the operators are invertible, these matrices are invertible as well. Doing so, we can interpret the matrix representation of  $\Delta_k$  as the Schur complement of the matrix

$$M_k^N = \begin{bmatrix} D_k^N & U_k^N \\ L_{k+1}^N & \Delta_{k+1}^N \end{bmatrix}. \quad (\text{C.29})$$

It is well known from linear algebra that the determinant of  $M_k^N$  satisfies

$$\det(M_k^N) = \det(\Delta_{k+1}^N) \det(\Delta_k^N). \quad (\text{C.30})$$

When both  $D_k$  and  $\Delta_{k+1}$  are invertible, the matrix  $M_k^N$  is invertible. Hence, note from (C.30) that, for  $k \geq 1$ , the matrix  $\Delta_k^N$  can be inverted for any  $N$ , and therefore  $\Delta_k$  (as an operator from  $\mathcal{F}$  to  $\mathcal{F}$ ) is invertible.

The case  $k=0$  requires special care. In this case  $D_0$  is not invertible and the previous argument cannot be applied. In order to make the solution unique, we need to impose an additional constraint to  $f^{(0)}$ . On ergodic flux-surfaces, condition (38) is sufficient to fix the value of  $f^{(0)}$ . However, this is not always the case when  $\iota$  is rational. Condition (38) fixes the value of  $f^{(0)}$  solely when the only functions that lie simultaneously at the kernels of  $D_0 = -\widehat{E}_\psi \langle B^2 \rangle^{-1} \mathbf{B} \times \nabla \psi \cdot \nabla$  and  $L_1 = \mathbf{b} \cdot \nabla$  are constants (flux-functions). If  $\widehat{E}_\psi \neq 0$ , this occurs for any  $\delta \neq -1/\iota$ . However, the case  $\delta = -1/\iota$  is unphysical as it would imply  $\sqrt{g} = 0$ . Hence, in practice, when  $\widehat{E}_\psi \neq 0$  condition (38) is sufficient to fix the value of  $f^{(0)}$  even if the surface is not ergodic. For rational flux-surfaces and  $\widehat{E}_\psi = 0$ , condition (38) is insufficient to fix  $f^{(0)}$ . In such case, we would need to fix the value of  $f^{(0)}$  at a point of each field line as any function  $g(\alpha)$  lies in the kernel of  $\mathbf{b} \cdot \nabla$ . In order to clarify this assertion, let us try to obtain  $f^{(0)}$  assuming that  $f^{(1)}$  is

known. Integrating the Legendre mode  $k=1$  of equation (41) along a field line gives

$$f^{(0)}(\alpha, l) = f_0^{(0)}(\alpha) - \int_0^l (\sigma^{(1)} - \Delta_1 f^{(1)}) dl'. \quad (\text{C.31})$$

If  $\iota$  is irrational  $f_0^{(0)}$  does not depend on  $\alpha$ . In this case, equation (C.31) and condition (38) fix  $f^{(0)}$  for each  $\sigma^{(1)}, f^{(1)}$ . When  $\iota$  is rational we need to distinguish between the case with and without radial electric field.

- (i) For  $E_\psi = 0$ , the constant  $f_0^{(0)}$  is free as no other equation includes  $f^{(0)}$ . As  $f_0^{(0)}$  depends on  $\alpha$ , condition (38) does not fix this integration constant.
- (ii) For  $E_\psi \neq 0$ , inserting (C.31) in the Legendre mode  $k=0$  of equation (34) gives

$$\begin{aligned} -\frac{\widehat{E}_\psi}{\langle B^2 \rangle} B^2 \frac{\partial f_0^{(0)}}{\partial \alpha} &= s^{(0)} - U_0 f^{(1)} \\ -\frac{\widehat{E}_\psi}{\langle B^2 \rangle} B^2 \frac{\partial}{\partial \alpha} \int_0^l (\sigma^{(1)} - \Delta_1 f^{(1)}) dl' &. \end{aligned} \quad (\text{C.32})$$

Integrating  $\int_0^{L_c}$  equation (C.32)  $dl$  gives a differential equation in  $\alpha$  from which we can obtain  $f_0^{(0)}$  up to a constant. Thus, (C.31), condition (38) and (C.32) fix  $f^{(0)}$ .

Hence, in ergodic flux-surfaces or rational flux-surfaces with finite radial electric field,  $M_0^N$  has a one-dimensional kernel. Thus, for  $k=0$ , it is necessary to substitute one of the rows of  $[D_0^N \ U_0^N]$  by the condition (38) so that  $M_0^N$  is invertible for any  $N$  and as  $\Delta_1^N$  can be inverted, also  $\Delta_0^N$  constructed in this manner for any  $N$ , which implies that  $\Delta_0$  (as the limit  $\lim_{N \rightarrow \infty} \Delta_0^N$ ) is invertible.

## Appendix D. Fourier collocation method

In this appendix we describe the Fourier collocation (also called pseudospectral) method for discretizing the angles  $\theta$  and  $\zeta$ . This discretization will be used to obtain the matrices  $\mathbf{L}_k$ ,  $\mathbf{D}_k$  and  $\mathbf{U}_k$ . For convenience, we will use the complex version of the discretization method but for the discretization matrices we will just take their real part as the solutions to (18) are all real. We search for approximate solutions to equation (34) of the form

$$f^{(k)}(\theta, \zeta) = \sum_{n=-N_{\zeta 1}/2}^{N_{\zeta 2}/2-1} \sum_{m=-N_{\theta 1}/2}^{N_{\theta 2}/2-1} \tilde{f}_{mn}^{(k)} e^{i(m\theta+nN_p\zeta)} \quad (\text{D.1})$$

where  $N_{\theta 1} = N_\theta - N_\theta \bmod 2$ ,  $N_{\theta 2} = N_\theta + N_\theta \bmod 2$ ,  $N_{\zeta 1} = N_\zeta - N_\zeta \bmod 2$ ,  $N_{\zeta 2} = N_\zeta + N_\zeta \bmod 2$  for some positive integers  $N_\theta, N_\zeta$ . The complex numbers

$$\tilde{f}_{mn}^{(k)} := \left\langle f^{(k)}, e^{i(m\theta+nN_p\zeta)} \right\rangle_{N_\theta N_\zeta} \left\| e^{i(m\theta+nN_p\zeta)} \right\|_{N_\theta N_\zeta}^{-2} \quad (\text{D.2})$$

are the discrete Fourier modes (also called discrete Fourier transform),

$$\langle f, g \rangle_{N_\theta N_\zeta} := \frac{1}{N_\theta N_\zeta} \sum_{j'=0}^{N_\zeta-1} \sum_{i'=0}^{N_\theta-1} f(\theta_{i'}, \zeta_{j'}) \overline{g(\theta_{i'}, \zeta_{j'})} \quad (\text{D.3})$$

is the discrete inner product associated to the equispaced grid points (49), (50),  $\|f\|_{N_\theta N_\zeta} := \sqrt{\langle f, f \rangle_{N_\theta N_\zeta}}$  its induced norm and  $\bar{z}$  denotes the complex conjugate of  $z$ . We denote by  $\mathcal{F}^{N_\theta N_\zeta}$  to the finite dimensional vector space (of dimension  $N_\theta N_\zeta$ ) comprising all the functions that can be written in the form of expansion (D.1).

The set of functions  $\{e^{i(m\theta+nN_p\zeta)}\} \subset \mathcal{F}^{N_\theta N_\zeta}$  forms an orthogonal basis for  $\mathcal{F}^{N_\theta N_\zeta}$  equipped with the discrete inner product (D.3). Namely,

$$\left\langle e^{i(m\theta+nN_p\zeta)}, e^{i(m'\theta+n'N_p\zeta)} \right\rangle_{N_\theta N_\zeta} \propto \delta_{mm'} \delta_{nn'} \quad (\text{D.4})$$

for  $-N_{\theta 1}/2 \leq m \leq N_{\theta 2}/2$  and  $-N_{\zeta 1}/2 \leq n \leq N_{\zeta 2}/2$ . Thus, for functions lying in  $\mathcal{F}^{N_\theta N_\zeta}$ , discrete expansions such as (D.1) coincide with their (finite) Fourier series. The discrete Fourier modes (D.2) are chosen so that the expansion (D.1) interpolates  $f^{(k)}$  at grid points. Hence, there is a vector space isomorphism between the space of discrete Fourier modes and  $f^{(k)}$  evaluated at the equispaced grid.

Combining equations (D.1)–(D.3) we can write our Fourier interpolant as

$$\begin{aligned} f^{(k)}(\theta, \zeta) &= \mathbf{I}(\theta, \zeta) \cdot \mathbf{f}^{(k)} \\ &= \sum_{j'=0}^{N_\zeta-1} \sum_{i'=0}^{N_\theta-1} I_{i'j'}(\theta, \zeta) f^{(k)}(\theta_{i'}, \zeta_{j'}), \end{aligned} \quad (\text{D.5})$$

where  $\mathbf{f}^{(k)} \in \mathbb{R}^{N_s}$  is the state vector containing  $f^{(k)}(\theta_{i'}, \zeta_{j'})$ . The entries of the vector  $\mathbf{I}(\theta, \zeta)$  are the functions  $I_{i'j'}(\theta, \zeta)$  given by,

$$I_{i'j'}(\theta, \zeta) = I_{i'}^\theta(\theta) I_{j'}^\zeta(\zeta), \quad (\text{D.6})$$

$$I_{i'}^\theta(\theta) = \frac{1}{N_\theta} \sum_{m=-N_{\theta 1}/2}^{N_{\theta 2}/2-1} e^{im(\theta-\theta_{i'})}, \quad (\text{D.7})$$

$$I_{j'}^\zeta(\zeta) = \frac{1}{N_\zeta} \sum_{n=-N_{\zeta 1}/2}^{N_{\zeta 2}/2-1} e^{N_p i n (\zeta - \zeta_{j'})}. \quad (\text{D.8})$$

Note that the interpolant is the only function in  $\mathcal{F}^{N_\theta N_\zeta}$  which interpolates the data at the grid points, as  $I_{i'}^\theta(\theta_i) = \delta_{ii'}$  and  $I_{j'}^\zeta(\zeta_j) = \delta_{jj'}$ .

Of course, our approximation (D.5) cannot (in general) be a solution to (34) at all points  $(\theta, \zeta) \in [0, 2\pi) \times [0, 2\pi/N_p]$ . Instead, we will force that the interpolant (D.5) solves equation (34) exactly at the equispaced grid points. Thanks to the vector space isomorphism (D.2) between  $\mathbf{f}^{(k)}$  and the discrete modes  $\tilde{f}_{mn}^{(k)}$  this is equivalent to matching the discrete Fourier modes of the left and right-hand-sides of equation (34).

Inserting the interpolant (D.5) in the left-hand side of equation (34) and evaluating the result at grid points gives

$$\begin{aligned} &\left( L_k f^{(k-1)} + D_k f^{(k)} + U_k f^{(k+1)} \right) \Big|_{(\theta_i, \zeta_j)} \\ &= \left( L_k \mathbf{I} \cdot \mathbf{f}^{(k-1)} + D_k \mathbf{I} \cdot \mathbf{f}^{(k)} + U_k \mathbf{I} \cdot \mathbf{f}^{(k+1)} \right) \Big|_{(\theta_i, \zeta_j)}. \end{aligned} \quad (\text{D.9})$$

Here,  $L_k \mathbf{I}(\theta_i, \zeta_j)$ ,  $D_k \mathbf{I}(\theta_i, \zeta_j)$  and  $U_k \mathbf{I}(\theta_i, \zeta_j)$  are respectively the rows of  $\mathbf{L}_k$ ,  $\mathbf{D}_k$  and  $\mathbf{U}_k$  associated to the grid point  $(\theta_i, \zeta_j)$ . We can relate them to the actual positions they will occupy in the matrices choosing an ordenation of rows and columns. We use the ordenation that relates respectively the row  $i_r$  and column  $i_c$  to the grid points  $(\theta_i, \zeta_j)$  and  $(\theta_{i'}, \zeta_{j'})$  as

$$i_r = 1 + i + j N_\theta, \quad (\text{D.10})$$

$$i_c = 1 + i' + j' N_\theta, \quad (\text{D.11})$$

for  $i, i' = 0, 1, \dots, N_\theta - 1$  and  $j, j' = 0, 1, \dots, N_\zeta - 1$ . With this ordenation, we define the elements of the row  $i_r$  and column  $i_c$  given by (D.10) and (D.11) of the matrices  $\mathbf{L}_k$ ,  $\mathbf{D}_k$  and  $\mathbf{U}_k$  to be

$$(\mathbf{L}_k)_{i_r i_c} = L_k I_{i'j'}(\theta_i, \zeta_j), \quad (\text{D.12})$$

$$(\mathbf{D}_k)_{i_r i_c} = D_k I_{i'j'}(\theta_i, \zeta_j), \quad (\text{D.13})$$

$$(\mathbf{U}_k)_{i_r i_c} = U_k I_{i'j'}(\theta_i, \zeta_j). \quad (\text{D.14})$$

Explicitly,

$$\begin{aligned} L_k I_{i'j'} \Big|_{(\theta_i, \zeta_j)} &= \frac{k}{2k-1} \left( \mathbf{b} \cdot \nabla I_{i'j'} \Big|_{(\theta_i, \zeta_j)} \right. \\ &\quad \left. + \frac{k-1}{2} \mathbf{b} \cdot \nabla \ln B \Big|_{(\theta_i, \zeta_j)} \delta_{ii'} \delta_{jj'} \right), \end{aligned} \quad (\text{D.15})$$

$$\begin{aligned} D_k I_{i'j'} \Big|_{(\theta_i, \zeta_j)} &= -\frac{\hat{E}_\psi}{\langle B^2 \rangle} \mathbf{B} \times \nabla \psi \cdot \nabla I_{i'j'} \Big|_{(\theta_i, \zeta_j)} \\ &\quad + \frac{k(k+1)}{2} \hat{\nu} \delta_{ii'} \delta_{jj'}, \end{aligned} \quad (\text{D.16})$$

$$\begin{aligned} U_k I_{i'j'} \Big|_{(\theta_i, \zeta_j)} &= \frac{k+1}{2k+3} \left( \mathbf{b} \cdot \nabla I_{i'j'} \Big|_{(\theta_i, \zeta_j)} \right. \\ &\quad \left. - \frac{k+2}{2} \mathbf{b} \cdot \nabla \ln B \Big|_{(\theta_i, \zeta_j)} \delta_{ii'} \delta_{jj'} \right), \end{aligned} \quad (\text{D.17})$$

where we have used expressions (31) and (32) to write

$$\begin{aligned} \mathbf{b} \cdot \nabla I_{i'j'} \Big|_{(\theta_i, \zeta_j)} &= \frac{B}{B_\zeta + \iota B_\theta} \Big|_{(\theta_i, \zeta_j)} \\ &\quad \times \left( \iota \delta_{jj'} \frac{dI_{i'}^\theta}{d\theta} \Big|_{\theta_i} + \delta_{ii'} \frac{dI_{j'}^\zeta}{d\zeta} \Big|_{\zeta_j} \right), \end{aligned} \quad (\text{D.18})$$

$$\begin{aligned} \mathbf{B} \times \nabla \psi \cdot \nabla I_{i'j'} \Big|_{(\theta_i, \zeta_j)} &= \frac{B^2}{B_\zeta + \iota B_\theta} \Big|_{(\theta_i, \zeta_j)} \\ &\quad \times \left( B_\zeta \delta_{jj'} \frac{dI_{i'}^\theta}{d\theta} \Big|_{\theta_i} - B_\theta \delta_{ii'} \frac{dI_{j'}^\zeta}{d\zeta} \Big|_{\zeta_j} \right). \end{aligned} \quad (\text{D.19})$$

**Table E1.** Normalization factors for DKES results.  $d\psi/dr$  in  $\text{T} \cdot \text{m}$ ,  $K_{11}$  in  $\text{T}^{-2} \cdot \text{m}^{-2}$  and  $K_{31}$  in  $\text{T}^{-1} \cdot \text{m}^{-1}$ .

| Configuration | $d\psi/dr$ | $K_{11}$ | $K_{31}$ |
|---------------|------------|----------|----------|
| W7X-EIM       | 0.5237     | 3.6462   | 1.9095   |
| W7X-KJM       | 0.5132     | 3.7969   | 1.9486   |
| CIEMAT-QI     | 0.4674     | 4.5774   | 2.1395   |

We remark that, for  $k = 0$ , the rows of  $\mathbf{D}_0$  and  $\mathbf{U}_0$  associated to the grid point  $(\theta_0, \zeta_0) = (0, 0)$ , are replaced by equation (38). Finally, each state vector  $\mathbf{f}^{(k)}$  for the Fourier interpolants contains the images  $f^{(k)}(\theta_i', \zeta_j')$  at the grid points, ordered according to (D.11).

## Appendix E. Convergence of monoenergetic coefficients calculated by DKES

The code DKES gives an approximation to the monoenergetic geometric coefficients as a semisum of two quantities  $\widehat{D}_{ij}^-$  and  $\widehat{D}_{ij}^+$  by solving a variational principle [36]. For each coefficient, the output of DKES consists on two quantities  $\widehat{D}_{ij}^\mp K_{ij}$ , where  $K_{ij}$  are the normalization factors

$$K_{ij} := \left( \frac{d\psi}{dr} \right)^{-2}, \quad i, j \in \{1, 2\}, \quad (\text{E.1})$$

$$K_{i3} := \left( \frac{d\psi}{dr} \right)^{-1}, \quad i \in \{1, 2\}, \quad (\text{E.2})$$

$$K_{3j} := \left( \frac{d\psi}{dr} \right)^{-1}, \quad j \in \{1, 2\}, \quad (\text{E.3})$$

$$K_{33} := 1, \quad (\text{E.4})$$

to change from the radial coordinate  $\psi$  to  $r$ . In table E1, the normalization factors for the configurations considered are listed.

Apart from the normalization factors, there is still a nuance left for the parallel conductivity coefficient: the code DKES computes this coefficient measured with respect to the one obtained by solving the Spitzer problem

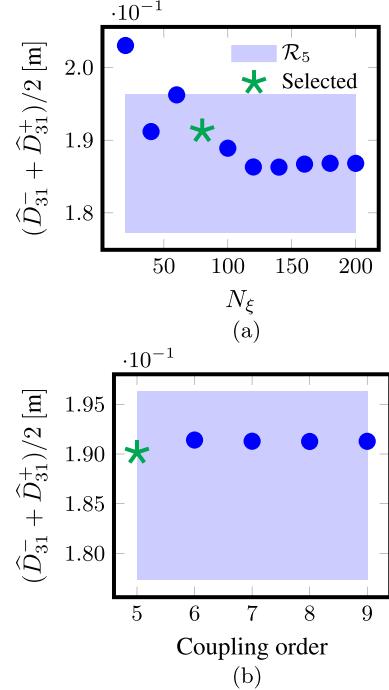
$$-\hat{\nu} \mathcal{L} f_{\text{Sp}} = s_3. \quad (\text{E.5})$$

Using (B.1) is immediate to obtain the 1-th Legendre mode of  $f_{\text{Sp}}$

$$f_{\text{Sp}}^{(1)} = \frac{1}{\hat{\nu}} \frac{B}{B_0} \quad (\text{E.6})$$

and using (47) we obtain its associated  $\widehat{D}_{33}$  coefficient

$$\widehat{D}_{33,\text{Sp}} = \frac{2}{3\hat{\nu}} \left\langle \frac{B^2}{B_0^2} \right\rangle. \quad (\text{E.7})$$



**Figure E1.** Convergence of  $(\widehat{D}_{31}^- + \widehat{D}_{31}^+)/2$  computed with DKES for W7X-EIM at the surface labelled by  $\psi/\psi_{\text{lefs}} = 0.200$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\widehat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 9. (b) Convergence with the coupling order for  $N_\xi = 80$ .

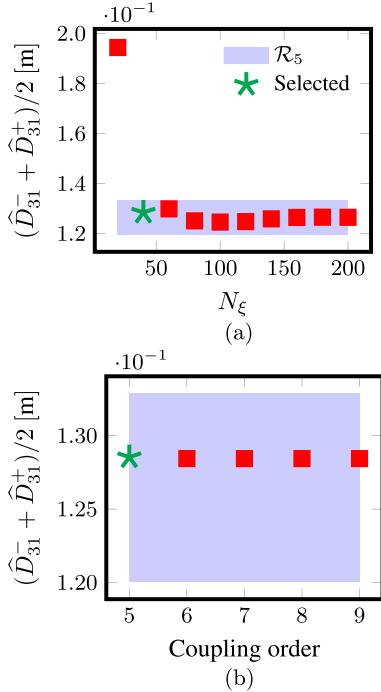
Thus, the output of DKES for the parallel conductivity coefficient has to be compared against the deviation  $(\widehat{D}_{33} - \widehat{D}_{33,\text{Sp}})$ .

From the output of DKES, the diagonal elements  $\widehat{D}_{ii}^\pm$  satisfy  $\widehat{D}_{ii}^- \geq \widehat{D}_{ii} \geq \widehat{D}_{ii}^+$  and allow to compute bounds for  $\widehat{D}_{ij}$

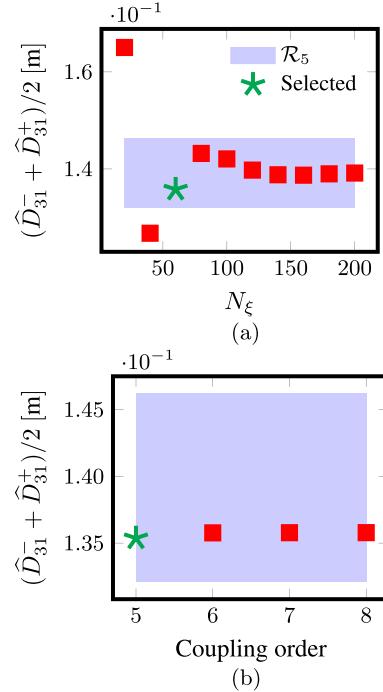
$$\frac{\widehat{D}_{ij}^- + \widehat{D}_{ij}^+}{2} - \Delta_{ij} \leq \widehat{D}_{ij} \leq \frac{\widehat{D}_{ij}^- + \widehat{D}_{ij}^+}{2} + \Delta_{ij} \quad (\text{E.8})$$

$$\text{and } \Delta_{ij} = \sqrt{(\widehat{D}_{ii}^- - \widehat{D}_{ii}^+)(\widehat{D}_{jj}^- - \widehat{D}_{jj}^+)/2}.$$

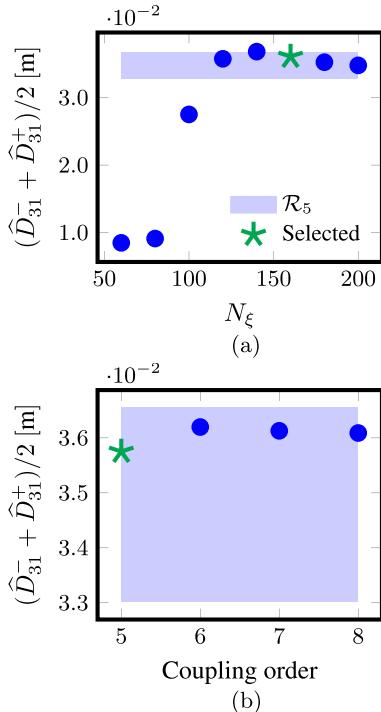
In figures E1–E6 the convergence study for selecting DKES resolutions is shown. In the code DKES the number of Legendre modes used are specified by  $N_\xi$ . In order to select the number of Fourier modes in the Boozer angles  $(\theta, \zeta)$  that DKES uses, an integer called ‘coupling order’ must be specified. Using figures E1(a), E2(a), E3(a), E4(a), E5(a) and E6(a), the number of Legendre modes  $N_\xi$  is selected so that it satisfies convergence condition (i) using the region  $\mathcal{R}_\epsilon$  for each case. After that, using E1(b), E2(b), E3(b), E4(b), E5(b) and E6(b), we select the minimum value of the coupling order for which the calculation with the selected value of  $N_\xi$  satisfies convergence condition (ii).



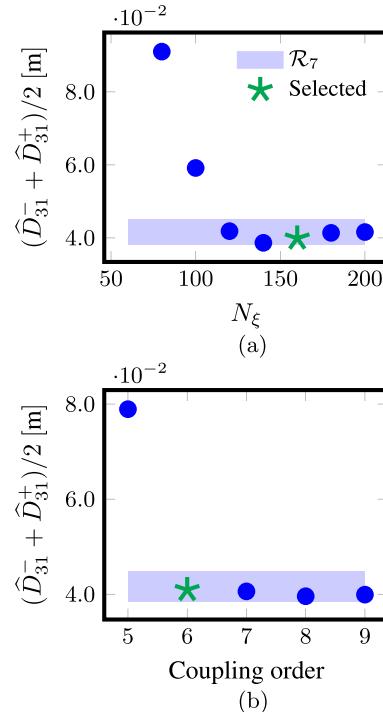
**Figure E2.** Convergence of  $(\hat{D}_{31}^- + \hat{D}_{31}^+)/2$  computed with DKES for W7X-EIM at the surface labelled by  $\psi/\psi_{\text{lcf}} = 0.200$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 3 \cdot 10^{-4} \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 9. (b) Convergence with the coupling order for  $N_\xi = 40$ .



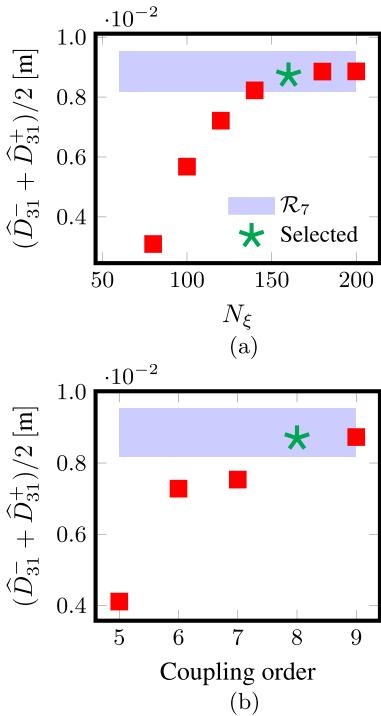
**Figure E4.** Convergence of  $(\hat{D}_{31}^- + \hat{D}_{31}^+)/2$  computed with DKES for W7X-KJM at the surface labelled by  $\psi/\psi_{\text{lcf}} = 0.204$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 3 \cdot 10^{-4} \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 7. (b) Convergence with the coupling order for  $N_\xi = 60$ .



**Figure E3.** Convergence of  $(\hat{D}_{31}^- + \hat{D}_{31}^+)/2$  computed with DKES for W7X-KJM at the surface labelled by  $\psi/\psi_{\text{lcf}} = 0.204$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 8. (b) Convergence with the coupling order for  $N_\xi = 160$ .



**Figure E5.** Convergence of  $(\hat{D}_{31}^- + \hat{D}_{31}^+)/2$  computed with DKES for CIEMAT-QI at the surface labelled by  $\psi/\psi_{\text{lcf}} = 0.250$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 0 \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 9. (b) Convergence with the coupling order for  $N_\xi = 160$ .



**Figure E6.** Convergence of  $(\hat{D}_{31}^- + \hat{D}_{31}^+)/2$  computed with DKES for CIEMAT-QI at the surface labelled by  $\psi/\psi_{lcf} = 0.250$ , for  $\hat{\nu}(v) = 10^{-5} \text{ m}^{-1}$  and  $\hat{E}_r(v) = 10^{-3} \text{ kV} \cdot \text{s m}^{-2}$ . (a) Convergence with  $N_\xi$  for coupling order = 9. (b) Convergence with the coupling order for  $N_\xi = 160$ .

## ORCID iDs

- F.J. Escoto <https://orcid.org/0000-0002-5315-7277>  
J.L. Velasco <https://orcid.org/0000-0001-8510-1422>  
I. Calvo <https://orcid.org/0000-0003-3118-3463>  
M. Landreman <https://orcid.org/0000-0002-7233-577X>  
F.I. Parra <https://orcid.org/0000-0001-9621-7404>

## References

- [1] Cary J.R. and Shasharina S.G. 1997 Omnigenity and quasihelicity in helical plasma confinement systems *Phys. Plasmas* **4** 3323–33
- [2] Helander P. and Nührenberg J. 2009 Bootstrap current and neoclassical transport in quasi-isodynamic stellarators *Plasma Phys. Control. Fusion* **51** 055004
- [3] Helander P., Geiger J. and Maaßberg H. 2011 On the bootstrap current in stellarators and tokamaks *Phys. Plasmas* **18** 092505
- [4] Beidler C.D. *et al* (The W7-X Team) 2021 Demonstration of reduced neoclassical energy transport in Wendelstein 7-X *Nature* **596** 221–6
- [5] Dinklage A. *et al* (The W7-X Team) 2018 Magnetic configuration effects on the Wendelstein 7-X stellarator *Nat. Phys.* **14** 855–60
- [6] Beidler C.D. *et al* 2011 Benchmarking of the mono-energetic transport coefficients—results from the International Collaboration on Neoclassical Transport in Stellarators (ICNTS) *Nucl. Fusion* **51** 076001
- [7] Sánchez E., Velasco J.L., Calvo I. and Mulas S. 2023 A quasi-isodynamic configuration with good confinement of fast ions at low plasma beta *Nucl. Fusion* **63** 066037
- [8] Velasco J.L., Calvo I., Sánchez E. and Parra F.I. 2023 Robust stellarator optimization via flat mirror magnetic fields (arXiv:2306.17506 [physics.plasm-ph])
- [9] Jorge R., Plunk G.G., Drevlak M., Landreman M., Lobsien J.-F., Camacho Mata K. and Helander P. 2022 A single-field-period quasi-isodynamic stellarator *J. Plasma Phys.* **88** 175880504
- [10] Camacho Mata K., Plunk G.G. and Jorge R. 2022 Direct construction of stellarator-symmetric quasi-isodynamic magnetic configurations *J. Plasma Phys.* **88** 905880503
- [11] Goodman A., Camacho Mata K., Henneberg S.A., Jorge R., Landreman M., Plunk G., Smith H., Mackenbach R. and Helander P. 2022 Constructing precisely quasi-isodynamic magnetic fields (arXiv:2211.09829 [physics.plasm-ph])
- [12] Pytte A. and Boozer A.H. 1981 Neoclassical transport in helically symmetric plasmas *Phys. Fluids* **24** 88–92
- [13] Boozer A.H. 1983 Transport and isomorphic equilibria *Phys. Fluids* **26** 496–9
- [14] Landreman M. and Paul E. 2022 Magnetic fields with precise quasisymmetry for plasma confinement *Phys. Rev. Lett.* **128** 035001
- [15] Helander P. 2014 Theory of plasma confinement in non-axisymmetric magnetic fields *Rep. Prog. Phys.* **77** 087001
- [16] Landreman M., Buller S. and Drevlak M. 2022 Optimization of quasi-symmetric stellarators with self-consistent bootstrap current and energetic particle confinement *Phys. Plasmas* **29** 082501
- [17] Anderson F.S.B., Almagri A.F., Anderson D.T., Matthews P.G., Talmadge J.N. and Leon Shohet J. 1995 The helically symmetric experiment, (HSX) goals, design and status *Fusion Technol.* **27** 273–7
- [18] Kinoshita S., Shimizu A., Okamura S., Isobe M., Xiong G., Liu H. and Xu Y. (The CFQS Team) 2019 Engineering design of the Chinese first quasi-axisymmetric stellarator (CFQS) *Plasma Fusion Res.* **14** 3405097
- [19] Zarnstorff M.C. *et al* 2001 Physics of the compact advanced stellarator NCSX *Plasma Phys. Control. Fusion* **43** A237
- [20] Nemov V.V., Kasilov S.V., Kernbichler W. and Heyn M.F. 1999 Evaluation of  $1/v$  neoclassical transport in stellarators *Phys. Plasmas* **6** 4622–32
- [21] Shaing K.C. and Callen J.D. 1983 Neoclassical flows and transport in nonaxisymmetric toroidal plasmas *Phys. Fluids* **26** 3315–26
- [22] Nakajima N., Okamoto M., Todoroki J., Nakamura Y. and Wakatani M. 1988 Optimization of bootstrap current in a large helical system with  $L = 2$  (available at: <https://ui.adsabs.harvard.edu/abs/1988obcl.rept.....N>)
- [23] Helander P., Parra F.I. and Newton S.L. 2017 Stellarator bootstrap current and plasma flow velocity at low collisionality *J. Plasma Phys.* **83** 905830206
- [24] Calvo I., Parra F.I., Velasco J.L. and Alonso J.A. 2017 The effect of tangential drifts on neoclassical transport in stellarators close to omnigenicity *Plasma Phys. Control. Fusion* **59** 055014
- [25] d'Herbemont V., Parra F.I., Calvo I. and Velasco J.L. 2022 Finite orbit width effects in large aspect ratio stellarators *J. Plasma Phys.* **88** 905880507
- [26] Velasco J.L., Calvo I., Parra F.I. and García-Regaña J.M. 2020 KNOSOS: a fast orbit-averaging neoclassical code for stellarator geometry *J. Comput. Phys.* **418** 109512
- [27] Velasco J.L., Calvo I., Parra F.I., d'Herbemont V., Smith H.M., Carralero D. and Estrada T. (The W7-X Team) 2021 Fast

- simulations for large aspect ratio stellarators with the neoclassical code KNOSOS *Nucl. Fusion* **61** 116013
- [28] Lazerson S., Schmitt J., Zhu C. and Breslau J. (STELLOPT Developers, All and USDOE Office of Science) 2020 STELLOPT, version 2.7.5 (available at: [www.osti.gov//servlets/purl/1617636](http://www.osti.gov//servlets/purl/1617636))
- [29] Bañón Navarro A., Di Siena A., Velasco J.L., Wilms F., Merlo G., Windisch T., LoDestro L.L., Parker J.B. and Jenko F. 2023 First-principles based plasma profile predictions for optimized stellarators *Nucl. Fusion* **63** 054003
- [30] Barnes M., Abel I.G., Dorland W., Görler T., Hammett G.W. and Jenko F. 2010 Direct multiscale coupling of a transport code to gyrokinetic turbulence codes *Phys. Plasmas* **17** 056109
- [31] Taguchi M. 1992 A method for calculating neoclassical transport coefficients with momentum conserving collision operator *Phys. Fluids B* **4** 3638–43
- [32] Sugama H. and Nishimura S. 2002 How to calculate the neoclassical viscosity, diffusion and current coefficients in general toroidal plasmas *Phys. Plasmas* **9** 4637–53
- [33] Sugama H. and Nishimura S. 2008 Moment-equation methods for calculating neoclassical transport coefficients in general toroidal plasmas *Phys. Plasmas* **15** 042502
- [34] Maaßberg H., Beidler C.D. and Turkin Y. 2009 Momentum correction techniques for neoclassical transport in stellarators *Phys. Plasmas* **16** 072504
- [35] Hirshman S.P., Shaing K.C., van Rij W.I., Beasley C.O. and Crume E.C. 1986 Plasma transport coefficients for nonsymmetric toroidal confinement systems *Phys. Fluids* **29** 2951–9
- [36] van Rij W.I. and Hirshman S.P. 1989 Variational bounds for transport coefficients in three-dimensional toroidal plasmas *Phys. Fluids B* **1** 563–9
- [37] Landreman M., Smith H.M., Mollén A. and Helander P. 2014 Comparison of particle trajectories and collision operators for collisional transport in nonaxisymmetric plasmas *Phys. Plasmas* **21** 042503
- [38] Helander P. and Sigmar D.J. 2005 *Collisional Transport in Magnetized Plasmas* (Cambridge University Press)
- [39] Landreman M. 2011 The monoenergetic approximation in stellarator neoclassical calculations *Plasma Phys. Control. Fusion* **53** 082003
- [40] Dewar R.L. and Hudson S.R. 1998 Stellarator symmetry *Physica D* **112** 275–80
- [41] Spong D.A. 2005 Generation and damping of neoclassical plasma flows in stellarators *Phys. Plasmas* **12** 056114
- [42] Trefethen L.N. and Weideman J.A.C. 2014 The exponentially convergent trapezoidal rule *SIAM Rev.* **56** 385–458
- [43] Anderson E. et al 1999 *Lapack Users' Guide* 3rd edn (Society for Industrial and Applied Mathematics)
- [44] Tennyson J.L., Cary J.R. and Escande D.F. 1986 Change of the adiabatic invariant due to separatrix crossing *Phys. Rev. Lett.* **56** 2117–20